

**Advanced Systems Engineering
and
Network Planning Support**

Deliverable: Final Report

**Contract No. NAS5-30680
Task 33**

submitted by

**CTA Incorporated
6116 Executive Boulevard
Rockville, MD 20852**

December 21, 1990

**(NASA-CR-189274) ADVANCED SYSTEMS
ENGINEERING AND NETWORK PLANNING SUPPORT
Final Report, 20 Aug. - 21 Dec. 1990
(Computer Technology Associates) 143 p**

N92-27972

**Unclas
G3/81 0097819**

Table of Contents

	Executive Summary	ES-1
1.	Introduction	1-1
	1.1 Objectives	1-1
	1.2 Scope	1-1
	1.3 Constraints	1-3
	1.4 Organization	1-3
2.	Requirements	2-1
	2.1 Overview	2-1
	2.2 User Categorization	2-1
	2.3 Modes of Operation	2-2
	2.4 Service Taxonomy	2-3
	2.5 Functional Analysis	2-4
3.	Architectural Alternatives	3-1
	3.1 Approach	3-1
	3.2 Resource Allocation	3-3
	3.2.1 Overview	3-3
	3.2.2 Pre-Planned	3-3
	3.2.3 Demand Access	3-11
	3.2.4 Short-Term Scheduling	3-14
	3.2.5 Resource Partitioning	3-15
	3.2.6 Hybrid Access	3-16
	3.3 Architectural Alternatives	3-17
	3.3.1 Overview	3-17
	3.3.2 SNC/ISC Functionality	3-19
	3.3.3 Architecture 1 - Current approach	3-21
	3.3.4 Real-Time vs Non Real-Time Functions	3-25
	3.3.5 Classified vs Unclassified Subsystems	3-25
	3.3.6 Standalone ISC Systems	3-29
	3.3.7 Integrated with SNC	3-31
	3.4 Infrastructure Support	3-32
	3.4.1 Introduction	3-32
	3.4.2 Distributed Data Management	3-32
	3.4.3 Communications	3-34
	3.4.4 Automated Interface Management	3-35
	3.4.5 Information Resource Dictionary	3-35
	3.4.6 Computer Security	3-37

Table of Contents (continued)

4.	Tradeoffs	4-1
4.1	Definition of Evaluation Criteria	4-1
4.2	Capability	4-1
4.2.1	Resource Allocation	4-2
4.2.2	Resource Partitions	4-3
4.2.3	Real-time vs Non-real-time Functional Partitions	4-3
4.2.4	Integration of SNC with ATGT	4-3
4.2.5	ISC	4-3
4.3	Relative Cost Analysis and Rankings	4-4
4.3.1	Resource Allocation	4-4
4.3.2	Resource Partitions	4-5
4.3.3	Real-time and non-real-time Partitions	4-6
4.3.4	Integration of Real-time SNC with ATGT	4-7
4.3.5	ISC System	4-7
4.4	Risk	4-8
4.4.1	Resource Allocation	4-9
4.4.2	Resource Partitions	4-9
4.4.3	Real-time vs Non-real-time Functional Partitions	4-10
4.4.4	Integrated SNC with ATGT	4-10
4.4.5	ISC	4-10
4.5	Performance	4-10
4.5.1	Resource Allocation	4-10
4.5.2	Resource Partitions	4-11
4.6	Ease of Use	4-11
4.6.1	Resource Allocation	4-11
4.6.2	Resource Partitions	4-12
4.6.3	Real-time vs. Non-real-time Partition	4-12
4.6.4	Integration of SNC with ATGT	4-12
4.6.5	ISC System	4-13
4.7	Flexibility	4-13
4.7.1	Resource Allocation	4-13
4.7.2	Resource Partitions	4-13
4.7.3	Real-time vs Non-real-time Functional Partition	4-13
4.7.4	Integration of SNC with ATGT	4-13
4.7.5	ISC System	4-13
4.8	Expandability	4-14
4.8.1	Resource Allocation	4-14
4.8.2	Resource Partitioning	4-14
4.8.3	Real-time Function Partitioning	4-14
4.8.4	Integration of Real-time Control System and ATGT	4-15
4.8.5	Inter-system Control (ISC)	4-15

Table of Contents (continued)

5.	Conclusions, Recommendations, and Rationale	5-1
5.1	Real-Time vs. Pre-Planned Resource Allocation	5-3
5.2	Information Interface	5-4
5.3	Control	5-5
5.4	Automation	5-6
5.5	Residency	5-7
5.6	Absorption of SNC by Other Systems	5-8
5.7	Systems Impact	5-10
	5.7.1 Infrastructure Support	5-10
	5.7.2 Other NASA Systems	5-10

Figures

Figure 1-1	System of Systems Context	1-2
Figure 1-2	Methodology	1-5
Figure 2.5-1	SN Functional Categorization	2-9
Figure 3.1-1	Formulation of Alternatives	3-2
Figure 3.2-1	Universe of Resource Allocation Algorithms	3-4
Figure 3.2.2-1	Elements of Scheduling Cost	3-6
Figure 3.2.2.4-1	Example Schedule Request	3-11
Figure 3.2.2.4-2	Example Flexible Schedule	3-12
Figure 3.2.3-1	Demand Access Functional Flow	3-15
Figure 3.3-1	Summary Resource Allocation Alternatives	3-24
Figure 3.3-2	Inter-System Control - Logical View	3-28
Figure 3.3-3	Architecture 1 Current Approach	3-29
Figure 3.3-4	Architecture 2 Real-time Function Partition	3-30
Figure 3.3-5	Architecture 3 Dual Subnet Approach	3-33
Figure 3.3-6	Architecture 4 Stand Alone ISC	3-30
Figure 3.4-1	Current Approach	3-39
Figure 3.4-2	Recommended Approach	3-40
Figure 3.4-3	Secure Network Topology	3-44
Figure 3.4.6-1	Multi-level System Architecture	3-46
Figure 3.4.6-2	Functionally Partitioned Architecture	3-47
Figure 3.4.6-3	Data Partitioned Architecture	3-48

Table of Contents (continued)

Tables

Table 2.5-1	Level 1 and Level 2 Functional Decomposition	2-5
Table 3.3-1	SNC/ISC Functions	3-19
Table 3.3-2	Example Interface MIB	3-24
Table 3.3-3	Real-time SNC/ISC Functions	3-26

Appendix

Appendix A	A-1
Appendix B	B-1
Appendix C	C-1
Appendix D	D-1
Appendix E	E-1
Appendix F	F-1

Acronyms

ADP	Automatic Data Processing
ADPE	Automatic Data Processing Equipment
AOS	Acquisition of Signal
ATDRSS	Advanced Tracking and Delay Relay Satellite System
ATGT	Advanced TDRSS Ground Terminal
CCITT	Consultive Committee for International Telegraph and Telephone
CCSDS	Consultive Committee for Space Data Systems
CDOS	Customer Data Operations System
CIF	Communication Interface Functions
COBE	Cosmic Background Explorer
COMS	CDOS Operations Management Services
COTS	Commercial Off The Shelf
CRC	Cyclic Redundancy Check
CTA	CTA INCORPORATED
DDS	Data Delivery Services
DIS	Data Delivery Service
DSN	Deep Space Network
DoD	Department of Defense
DWM	Distribution Work Management
EOS	Earth Observing System
ERBS	Earth Radiation Budget Satellite
ESA	European Space Agency
FERN	Flexible Envelope Request Notation
FDF	Flight Dynamics Facility
GN	Ground Network
GOSIP	Government OSI Profile
GSFC	Goddard Space Flight Center
H&S	Health & Safety
IP	Internet Protocol
IRDS	Information Resource Dictionary System
ISC	Inter-System Control
ISCC	Inter-System Control Center
ISO	International Standards Organization
JPL	Jet Propulsion Lab
JSC	Johnson Space Center
KSAR	K-Band SAR

LAN	Local Area Network
LR	Lead Ranges
MA	Multi-access
MIB	Management Information Bases
MLS	Multi-level Secure
MPT	Mission Planning Terminal
MSOCC	Multisatellite Operations Control Center
NASA	National Aeronautics and Space Administration
NASCOM	NASA Communications Network
NASCOM II	The Next Generation Communications Network
NASDA	Japanese National Space Agency
NCC	Network Control Center
NGT	NASA Ground Terminal
NIST	National Institute of Standards and Technology
NMS	Network Management System
NPAS	Network Planning and Analysis System
OPNET	Optimized Network Engineering Tools
OSI	Open Systems Interconnection
PING	Packet Internet Groper
POCC	Payload Operations Control Center
QOS	Quality of Service
SA	Single Access
SAF	Single Access Forward
SAR	Single Access Return
SDNS	Secure Data Network System
SDPF	Sensor Data Processing Facilities
SNC	Space Network Control
SN	Space Network
SNIF	SN Interface Function
SSAR	S-Band SAR
STDN	Space Tracking and Data Network
STGT	Second TDRSS Ground Terminal
STS	Space Transportation System
TBD	To Be Determined
TDRSS	Tracking and Data Relay Satellite System
TOCC	TDRS Operations Control Center
USAT	User Satellites
WAN	Wide Area Network
WSC	White Sands Complex

Executive Summary

The objective of this task was to take a fresh look at the NASA Space Network Control (SNC) element for the Advanced Tracking and Data Relay Satellite System (ATDRSS) such that it can be made more efficient and responsive to the user by introducing new concepts and technologies appropriate for the 1997 time frame. In particular, it was desired to investigate the technologies and concepts employed in similar systems that may be applicable to the SNC. It is intended that these results will be used as input to the Phase A studies for the SNC.

This study was a short, intensive effort that focused on a small number of key issues. As the initial activity in this task, a high-level analysis of requirements and operation of the existing Network Control Center (NCC) was performed. This analysis established a baseline set of requirements for the SNC but also provided additional insight into the key issues. The key issues with some additional elaboration are summarized as follows:

Key Issue 1: What processing/scheduling is done in real-time versus pre-planned?

Elaboration: The current scheduling process is lengthy with many changes occurring. Although service can be provided on short notice, few users take advantage of this capability.

Key Issue 2: What is the system "information interface" for operators and users?

Elaboration: There is limited information provided back to the user during the scheduling process because the composite schedule is classified. This makes conflict resolution difficult. Also, the scheduling of the shuttle is the major source of perturbations whose impact affects many users.

Key Issue 3: What processing is automated versus manual?

Elaboration: Both the operation and use of the system could be made more efficient and less manually intensive. Increasing the level of automation of the existing NCC is expensive because modification of its software is difficult.

Key Issue 4: How is the system controlled (centralized/distributed)?

Elaboration: Rather than addressing just centralized or distributed control, a broader set of alternatives had to be considered involving hybrid approaches. Furthermore, control was addressed in terms of the individual OSI network management functions (configuration, fault, performance, security, and accounting).

Key Issue 5: Where is the processing performed? Where is the data stored?

Elaboration: The primary alternatives to be considered are the White Sands Complex (WSC), Goddard Space Flight Center (GSFC), and User Payload Operations Control Centers (POCCs).

Key Issue 6: Can the SNC be absorbed by other systems?

Elaboration: A substantial number of the SNC functions will be performed by the Second TDRSS Ground Terminal (STGT). The primary option is to migrate additional functionality to the Advanced TDRSS Ground Terminal (ATGT). To a lesser extent functions could be moved to the user POCCs, CDOS, or NASCOM II.

In order to identify new concepts and to resolve these issues, the following categories of systems were surveyed:

- o systems with similar space network control missions for satellite data collection,
- o commercial satellite and telecommunications common carriers,
- o systems whose resource allocation and control functions are functionally analogous and are referred to as "abstract analogues."

The first two types of systems were surveyed by site visits while the "abstract analogues" were surveyed by performing a literature search. The conclusion of this survey was that there is no other system that is a one-to-one matching to the SNC functions, but several useful concepts can be adopted for the SNC.

The major recommendations derived from this study address one or more of the key issues. These recommendations are listed below with a cross reference to the relevant key issue(s) in parentheses.

1. Adopt a Hybrid Scheduling Approach to Reduce Scheduling Operational Complexity and Achieve Maximum User Satisfaction (#1)

Although from a functional view, the SN could be operated like the telephone network, this is not practical because there is not enough SA service capacity. A much larger fraction of user service requests can be satisfied with the use of a pre-planned approach because there is time to resolve conflicts. However, demand access can be provided to support the needs of some users on the MA service with sufficiently low blocking. Therefore, a hybrid scheduling scheme is proposed. This would enable the SN to 1.) accommodate unforeseen needs as they occur in real-time as well as requirements that are known well in advance and 2.) reduce the scheduling workload.

For the pre-planned users, it is recommended to investigate a fluid scheduling approach with a shorter scheduling time horizon. Emulating the "just in time" concept of job shop scheduling, this approach would establish multiple "freeze points" at which time resources would be allocated. Since resources wouldn't be prematurely allocated, the impact of changes would be minimized. However, conflicts would still be resolved as they are identified.

One of the key results is the impact of the user providing a flexible request. Our simulations show that if the users provide a window time specification rather than a fixed time epoch, blocking is substantially reduced. The major issue to resolve is when the window must be converted to a fixed time assignment. The SNC could perform this scheduling in near real-time with adequate processing power, but users probably need more time (hours to days) to generate/fine tune their command set. Thus, the time horizon for scheduling is dominated by user constraints.

Of the satellite data collection facilities surveyed, both were utilizing a shorter scheduling horizon of 3 to 7 days. Adoption of a shorter horizon may reduce the impact of changes, but may not be feasible for all users. The use of a fluid scheduling approach with multiple freeze points will allow the planning horizon to be application specific such that the varying needs of the users can be supported.

2. Incorporate Resource Partitions to Isolate Impact of Users (#1, #2)

It is recommended that the capability of establishing resource partitions for subnetworks be incorporated into the SNC requirements. This will provide the capability to isolate the impact of various classes of users from each other. If the user requirements vary significantly by time of day, these partitions could be time variant similar to the routing employed in the telephone network.

The use of resource partitions is a way to reduce the impact of manned space flight. For example, SA channels could be assigned to the shuttle. However, in order to efficiently use these resources, a standby schedule would be needed in case of launch delays. Also, if after launch the shuttle schedule was available for on-line access, users might be able to utilize the remaining service times more easily.

Resource partitioning also introduces the possibility of providing on-line access to a schedule for some subset of the SN resources similar to the GTE video scheduling system. Access to the schedule would eliminate the need to send reject notices in response to a schedule request without any explanation, one of the major frustrations with the current NCC. Furthermore, this could allow the users to at least partially resolve their conflicts prior to submission of their requests, and thus reducing the workload on the SNC.

The downside of resource partitioning is the negative impact on performance. This is especially critical in case of SA channel failure. However, the simulation results indicate that there is potentially adequate capacity to partition the MA resources, without introducing a performance problem.

3. Further Automate the Entry, Change, and Conflict Resolution of Schedule Data (#1, #2, #3)

In the scheduling area, the handling of conflicts by voice co-ordination would be largely replaced in an incremental fashion with the semi-automated generation of shift requests, distributed data management to concurrently update schedules at the SNC and POCCs, and distributed work management tools to co-ordinate the group execution of the shift requests by people.

There is substantial potential for the automated generation of shift requests including, ultimately, the application of co-operating expert systems. However, there will probably always be some aspect of this activity involving manual intervention. This will require scheduling decision aids to support the analyst. For example, a tool was demonstrated by the Air Force that displayed a window of the schedule on a large monitor, provided the capability to enter changes to the contact time graphically, and identified conflicts.

The distributed data management and distributed work management technologies will be provided by off-the-shelf technologies. They are currently available in vendor products and will be further supported by Open Systems Interconnection (OSI) communications standards in the time frame for SNC implementation.

4. Automate the Inter-System Control Function and Monitor by Exception (#3, #4)

Automation of the Inter-System Control function is recommended with the monitoring and analysis of the network being performed on an exception basis such that operators are only notified when a problem is detected. In this concept an operator would be assigned to each contact to ensure that the quality of SN service is maintained, but the operator will be supporting multiple contacts concurrently. In order to do this, the operator must be able to obtain the "big picture" of the real-time system status on demand.

The automation of the ISC is based on the OSI concepts of a manager, agents resident in the systems being coordinated, and Management Information Base (MIB). The capabilities envisioned to be automated are pre-pass testing by the ISC agents, reporting of summary status by the ISC agents to the ISC manager, analysis of the status data by the ISC manager, generation of alerts by the ISC manager, and display of the system status by the ISC manager. The MIB structure defines the objects being managed for each system being coordinated. At the current time, the standardization process has only defined MIBs for components rather than systems. Also, the development of the systems being coordinated is leading the SNC development. Thus, this is a risk area that needs near-term attention.

Monitoring by exception is a major change in the operations concept from the existing NCC concept as operators will no longer be assigned to each pass. Although all of the satellite data collection facilities surveyed had an operator watching each pass, monitoring by exception is achievable with current technology. Since it has not been done, there is a risk involved. The first concern is the user acceptance of this approach; locating the ISC operators at GSFC may facilitate this acceptance. Second, it is imperative that the infrastructure be introduced to support the distributed data management of the ATDRSS configuration so that the SNC and the POCC have the same configuration; operators will not have the time to sort out these parameters under the new concept. Third, the handling of the perturbations introduced by the shuttle will have to

be streamlined. This can be done by allowing the users on-line access to a selected subset of the SNC schedule and using a standby schedule.

5. Implement a Real-time / Non-Real time Partition of SNC Subsystem Functionality (#4, #5, #6)

It is recommended that the SNC functions be partitioned into real-time and non-real-time subsystems with the real-time subsystem being integrated with ATGT and the non-real-time subsystem resident at GSFC in order to minimize cost as well as to achieve the most robust performance.

The non-real time subsystem, primarily performing the resource allocation function, would be located at GSFC. This is recommended because the schedule processing functions are significantly different (transaction processing vs. real-time communications) from the STGT functions and are not prime candidates for integration into ATGT. Furthermore, the primary traffic flows would be between the user POCCs at GSFC and the non-real-time subsystem. Although the traffic flows are not large compared to science data, the users will not be affected by congestion or failures in NASCOM when accessing the non-real-time system (if it is located at GSFC). This is especially important with the recommended increase in the level of SNC automation and data communications. Therefore, the non-real-time subsystem should be located closer to the users at GSFC.

From a functional point of view, it is reasonable to integrate the real-time control functions into the ATGT because of the similarity to the functions currently performed. The principal functions that could be integrated are the validation of POCC commands to modify the ATDRSS channel configuration during a support and the handling of demand access requests. Since the existing ground terminal already validates commands, this function should be integrated into ATGT. The tradeoffs associated with the integration of the demand access function are more complex and depend on specific functionality. First, a single point of processing in ATGT would have to be established to allocate resources so the user would not have to know which ground terminal to access; this requires an upgrade to the STGT architecture. Second, if demand access is a simple function providing service on a FIFO basis, then it is beneficial to integrate it into the ATGT. However, this function will be more complex if queueing of demand access requests is performed, and the real-time SNC performs some "look ahead" processing in order to optimally allocate SN resources. In this case, it is less beneficial to integrate demand access into ATGT.

Another major issue to be considered in this analysis is the upgrading of the STGT security functionality. If the real-time SNC functions are integrated into ATGT, then ATGT will communicate directly with unclassified POCCs in a transaction mode. Since STGT does not have this capability, its security architecture will have to be upgraded with the introduction of a Restricted Access Processor. This introduces some risk.

In summary, although integrating the real-time SNC functions into ATGT requires a further software complexity analysis as well as a timing and sizing analysis, this is the most attractive candidate for integration.

6. Introduce Automated Interface Management (#3, #4)

It is envisioned that the communications interface software between SNC components and external systems will be largely off-the-shelf OSI based software components. In the OSI environment, application message definitions are specified in a programming language representation, referred to as the Abstract Syntax Notation One (ASN.1). This will enable the use of ASN.1 compilers, an off-the-shelf OSI utility, to generate new encoding/decoding software when interfaces are modified. This will make changing interfaces simple and efficient, which may be especially important for accommodating the international partners as new requirements occur.

In summary, the results of this study include recommendations for the introduction of new concepts and technologies. These recommendations include resource partitioning, on-line access to subsets of the SN schedule, fluid scheduling, increased use of demand access on the MA service, automating Inter-System Control functions using OSI concepts and monitor by exception, increased automation for distributed data management and distributed work management, viewing SN operational control in terms of the OSI Management framework, and the introduction of automated interface management.

SNC Study Final Report

1. Introduction

1.1 Objectives

The objective of this task is to take a fresh look at the National Aeronautics and Space Administration (NASA) Space Network Control (SNC) element for the Advanced Tracking and Delay Relay Satellite System (ATDRSS) to make it more efficient and responsive to the user by introducing new concepts and technologies appropriate for the 1997 time frame. In particular, it is desired to investigate the technologies and concepts employed in similar systems that may be applicable to the SNC. These similar systems include:

- o systems with a similar space network control missions, e.g., Department of Defense(DoD) systems,
- o commercial satellite and telecommunications common carriers,
- o systems whose resource allocation and control functions are functionally analogous.

Because this study was performed in a short time period, it was necessary to focus it on a small number of key issues. Thus, the following key issues were identified at the outset of this study:

- 1. What processing/scheduling is done in real-time versus pre-planned?**
- 2. What is the user-system "information interface"? What is the operator user-system "information interface"?**
- 3. What processing is automated versus manual?**
 - o user services
 - o operator services
- 4. How is the system controlled (centralized/distributed)?**
- 5. Where is the processing performed? Where is the data stored?**
- 6. Can the SNC be absorbed by other systems?**

The results of this study will be a set of recommendations regarding these issues for input to Phase A SNC studies to be performed in 1991. These recommendations involve specification of SNC requirements or formulation of concepts to be evaluated in the Phase A studies.

1.2 Scope

The Space Network (SN) can be viewed as consisting of the SN communications satellite assets resident in space, the ground communications assets, and SNC element for the control of these assets. In addition, the SN must be viewed as a system within a "system of systems" because it must interface with a number of other systems. As shown in Figure 1-1, the SN is a service provider that must interoperate with other service providers and SN users. The SN consists of:

- o the ATDRSS,
- o the Advanced TDRSS Ground Terminal (ATGT),
- o Ground Network (GN) assets,
- o SNC for control of these assets.

The first three systems listed above are referred to as the communications assets because they perform the delivery of the user data. The functions of the SNC may be embedded in the communications assets, resident in stand-alone system(s) or partially embedded and partially stand-alone.

The other service providers within this "system of systems" context consist of both NASA and non-NASA systems. The NASA systems are:

- o the Jet Propulsion Laboratory (JPL) Deep Space Network of ground system,
- o Customer Data Operations System (CDOS) in the ATDRSS era,
- o Flight Dynamics Facility (FDF),
- o other existing sensor data processing facilities, and
- o NASA communication network (NASCOM) communications utility,

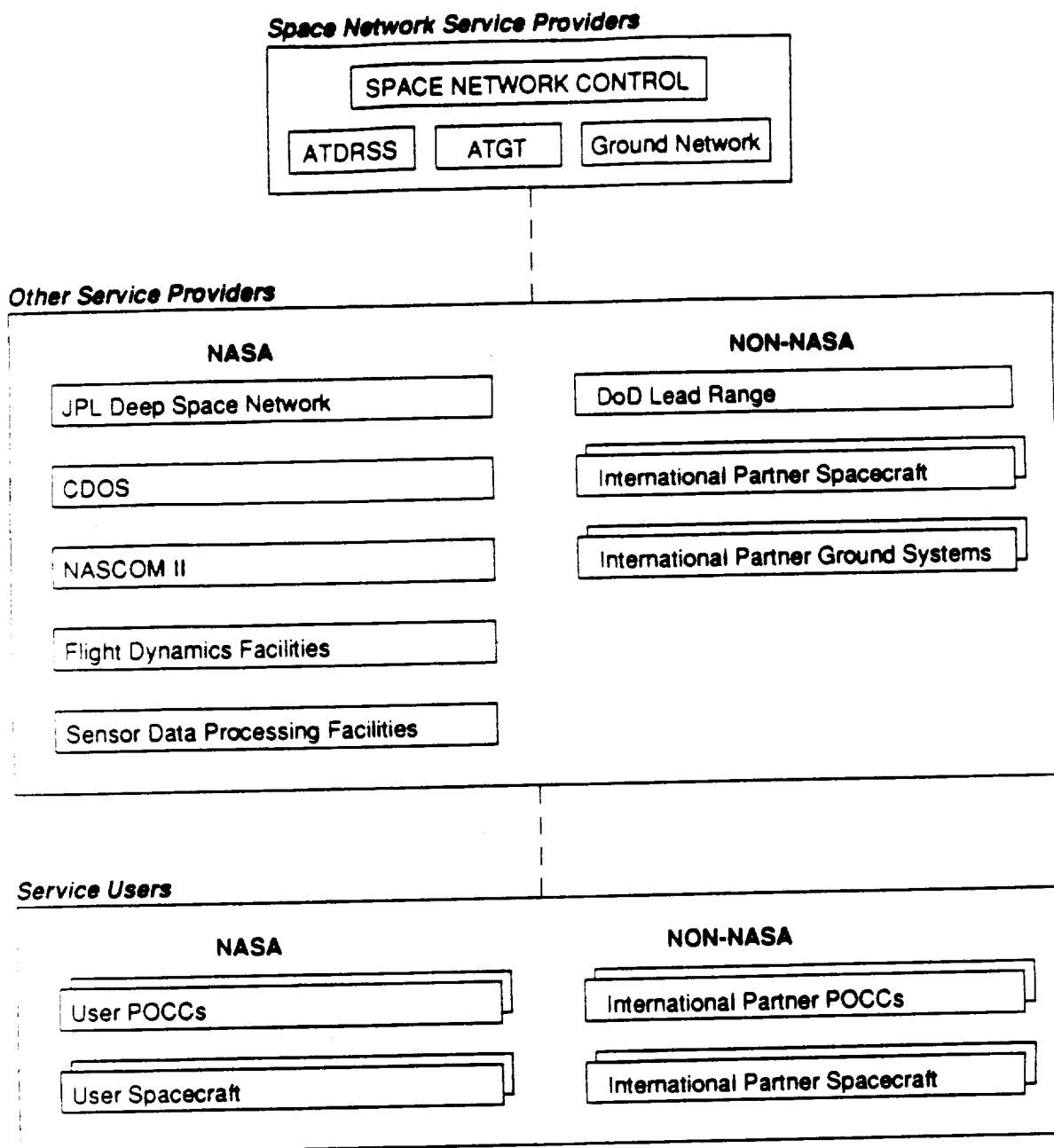


Figure 1-1. System of Systems Context

The non-NASA systems include the DoD Lead Range ground systems, as well as satellite systems/ground of the international partners [European Space Agency (ESA) and Japanese National Space Agency (NASDA) as either users or service providers.] The service user system consist of the user spacecraft and payload operations control centers (POCCs) for both NASA users as well as the international partners.

The areas that have most critical impact on the SNC and are managed by Code 500 are NASCOM II, CDOS, and the user POCCs. Thus, their interface with the SN is of particular importance in this study.

The scope of this work includes all aspects of SNC functionality. However, the major focus will be on the resource allocation of the communications assets and the assuring the performance of these assets rather than the administrative or sustaining engineering aspects of the SNC. It is assumed that the characteristics of the communications assets, e.g., number of channels/antennas, data rates, etc. are fixed so the focus of the work is on control of the communications assets. As identified, modifications or enhancements to the assets may be recommended.

This study was performed in parallel with on-going CTA INCORPORATED (CTA) studies on the SN and SNC. This study differs from these efforts in that:

- o surveys systems being used/developed by other organizations with analogous control problems,
- o is issue oriented as discussed above rather than comprehensive,
- o address the general system of systems issue rather than assume the existing allocations among systems.

Upon completion the results of this study will be integrated into these other ongoing efforts as appropriate.

1.3 Constraints

The three primary constraints affecting the SNC are its integration into the 1997 NASA environment, its interface with external systems, and security. In developing new approaches, it is recognized that they must be integrated into the existing and planned systems. As discussed above, this integration must be considered in a "system of systems" context. Any impact of the SNC on these other systems must be identified and evaluated.

In some cases users of the SN may want to utilize resources such as the DoD Lead Range facilities or the assets of the international partners. Thus, the users will have to co-ordinate the mode of operation for these systems, e.g., preplanned, demand access, or hybrid.

Security is a constraint because the system must be protected from unauthorized users. Also, the composite schedule showing the allocation of resources to users is currently classified. Thus, the protection of the system and confidentiality of data must be considered in developing new approaches.

1.4 Organization

In this section the overall approach for performing the work is presented and related to the organization of this report. The overall methodology used for performing this task is depicted in Figure 1-2. The initial efforts involved information gathering and consisted of requirements analysis, familiarization with the NASA Space Tracking and Data Network (STDN) environment, and survey of technology and similar systems. The results of the requirements analysis, as presented in Section 2, addressed the categorization of users, a taxonomy of services, and functional requirements. The set of functional requirements used in this study are presented in Appendix A.

The environment review element of the methodology involved familiarization with the status and plans of major NASA programs that will affect the SNC. These programs included:

- o ATDRSS program,
- o the existing TDRSS ground system at the White Sands Complex (WSC) and the upgraded ground system, referred to as the Second TDRSS Ground Terminal (STGT),

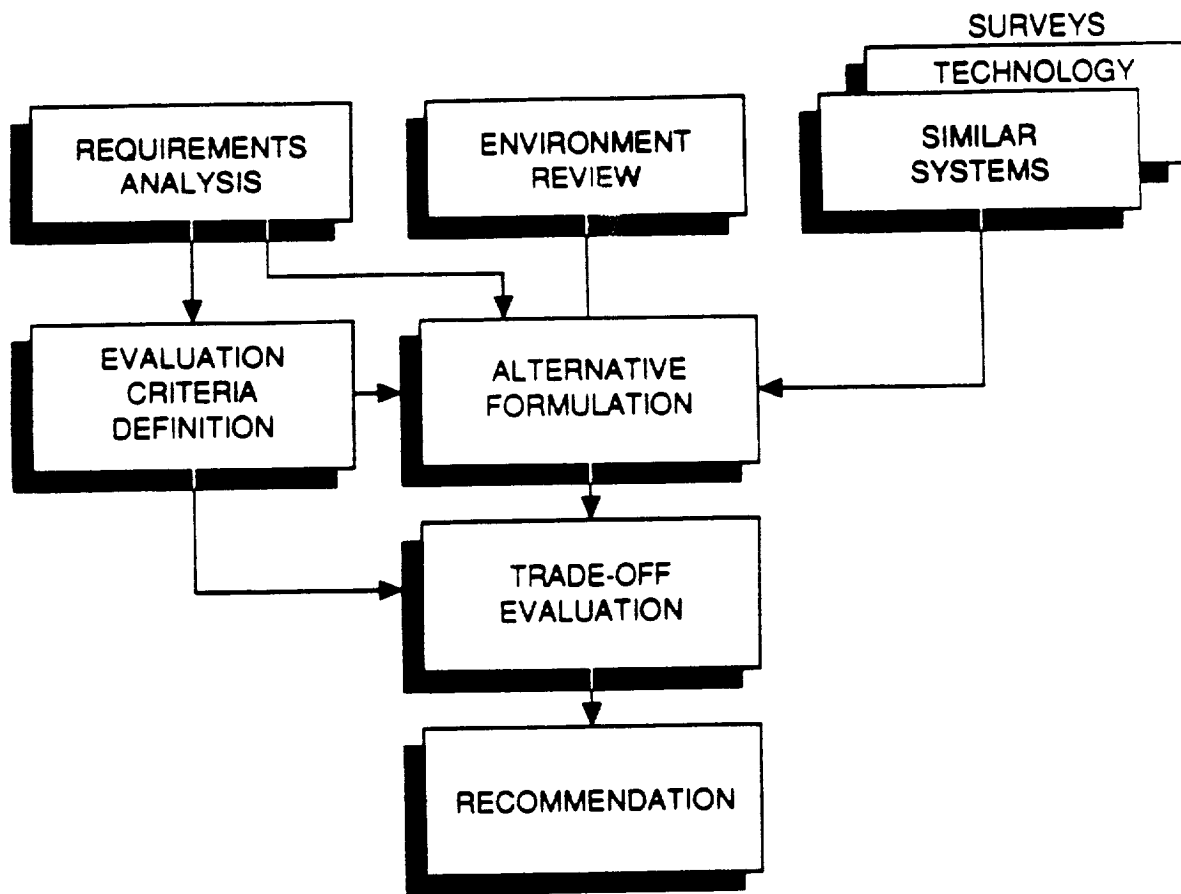


Figure 1-2. METHODOLOGY

- o the next generation communications network (NASCOM II),
- o the next generation data processing system (CDOS).

Also, the operation of the existing Network Control Center (NCC) was reviewed. These reviews were accomplished by document review, site visits, and briefings. Since this information is background data for performing the study, it is not specifically documented in this report, but appropriate references are provided where they affect the study results.

The technology survey was conducted by performing an electronic literature search in the areas of resource allocation and control. In this survey, the goal was to identify "abstract analogues," i.e., problems with a mathematical formulation similar to that of the SNC resource allocation and control problem. The results of this survey are summarized in Appendix B. The survey of similar systems addressed both systems being used or developed by other organizations with analogous resource allocation and control requirements. Organizations involved with satellite data collection as well as satellite communications were surveyed. The results of these site visits are presented in Appendix C. Although no one-to-one correlation was found between the SNC control and resource allocation problems with those surveyed, elements of the survey results were useful in the formulation of alternatives.

The next set of major steps involved the formulation and evaluation of architectural alternatives. The alternatives, as presented in Section 3 with supporting technical material in Appendix D, are formulated in terms of resource allocation and management. Management of the SNC as well as the inter-system management of the system of systems (SN, CDOS, NASCOM II) are addressed in this section. The criteria are defined and applied to these alternatives to identify the tradeoffs in Section 4. As part of the tradeoff evaluation, a performance analysis of alternative resource allocation schemes and a data flow analysis of alternative functional allocation schemes were performed. These results are presented in Appendices E and F, respectively.

Then the results of the tradeoffs were analyzed and a set of recommendations were prepared. These recommendations are presented in Section 5 and address each of the key issues identified at the outset of the study.

2. Requirements

2.1 Overview

In this section the results of the analysis of the high level Space Network Control (SNC) functional, operational, and performance requirements are presented. The initial activities involved preparing a categorization of SN users and a taxonomy of user services describing the potential services to be provided to the user. In preparing the user categorization, modes of operation, and service taxonomy, it was intended to cover all practical possibilities. These results, presented in Section 2.2 through 2.4, were useful in formulating new operational concepts for the Space Network (SN).

The primary sources of input for this activity were document review and site visits. Interviews were conducted with:

- o Space Transportation System (STS) (referred to as the Shuttle) users at Johnson Space Center (JSC),
- o Space Telescope users at Goddard and the Johns Hopkins Science Institute, and
- o Multi-mission users [Cosmic Background Explore (COBE), Earth Radiation Budget Satellite (ERBS)] at Goddard.

Space Station (SS) and Earth Orbiting System (EOS) requirements were also investigated.

In the requirements analysis, the functional definition that was prepared in the parallel "B team" SNC study was used as input to establish a baseline; it was reviewed and enhanced as necessary to address the key issues. The functional analysis is presented in Section 2.5.

In order to quantitatively analyze alternatives, a traffic model was formulated and parameters estimated as part of the performance requirements. This was done to a level of detail such that the feasibility, performance, and computational complexity of the scheduling algorithms could be understood. This data is key input for evaluating the computing capacity needed to implement various scheduling algorithms and the impact of changes in schedule requests on computing capacity. The traffic loading prepared for this study is summarized in Appendix E.

2.2 User Categorization

Tracking & Data Relay Satellite System (TDRSS) users can be divided into three major categories, spacecraft operations, mission / science users, and testing / training. These users differ in their responsibilities and objectives and therefore present a diversity of requirements for Network Control Center (NCC) and SNC. Spacecraft operations includes those personnel who are responsible for maintaining the orbit of the spacecraft, monitoring the status of its subsystems, and insuring the safety of the spacecraft and its instruments. Mission / science users are those individuals who use the spacecraft and (optionally) its instruments to meet their science or mission objectives. Testing / training users have yet another set of objectives, to train new personnel to operate the various parts of the system, and to perform testing of new software releases / hardware upgrades before bringing them on-line and testing of the current baseline to locate bugs and isolate faults. The types of variability in the ways in which each of these categories of users uses TDRSS is described below.

SPACECRAFT OPERATIONS: There are roughly four ways in which spacecraft operations users may use TDRSS. In some cases engineering and / or health and safety data will be written to tape and played back when the satellite is in view of TDRSS. In most cases the data of this type will be received in real-time. A further variation of the real-time case occurs when a real-time forward link for commanding is also required. These activities are normally conducted on a routine basis and are scheduled in advance; however, these users must have the ability to request emergency service if the satellite becomes unstable or at risk. Of course there will also always be some degree of routine changes in the service requirements of these users.

MISSION / SCIENCE: Mission / Science users of TDRSS present the most demanding requirements for the system. These users will have the same types of variability as the spacecraft operations personnel in that they will sometimes record, sometimes use a real-time return link to acquire their data, and sometimes require a real-time forward link in conjunction with a real-time return link. These users must also be able to obtain emergency service and

effect routine changes to their service requests. In addition some users' missions are of a short duration with uncertain launch times; other missions are long term with a large degree of stability. Some mission / science users need as nearly continuous coverage as possible; others need coverage only for parts of orbits or in some cases once every several orbits. Although most users use only one band at a time, STS Docking with another spacecraft requires S-Band Single Access Return (SSAR) and K-Band Single Access Return (KSAR) (on one TDRSS Single Access (SA) antenna) simultaneously. Some spacecraft are complex containing multiple instruments, instruments which can be used in several ways, or both. Other users' spacecraft incorporate a single instrument or which support a single or small number of investigators. The spacecraft or instruments which support some mission / science users required command loads which are very long (thousands). These users require an increased amount of advance notice of service to allow time for the command load to be finalized.

TESTING / TRAINING: Time must be blocked out of the schedule to support testing of new software releases and to train personnel. These activities can be fit into openings in the schedule and can be bumped when necessary. Time must also be taken from the schedule to support maintenance activities. Like other users, these users must be able to obtain emergency service and effect routine changes to their service requests. Emergency service may be required when failures occur.

2.3 Modes of Operation

There are several types of modes of mission operations, with differing needs for scheduling of SN resources. Most missions tend to operate in one mode, but at specific times must operate in another mode. How a conceptual alternative will respond to each of these modes is of prime importance in evaluating the alternative, especially in the case of the resource allocation.

Mode 1 - Single event operations

Single event operations include launch, orbit insertion, orbit maneuvers, and initial deployment. Operating in this mode requires guaranteed support, but start time is unpredictable, due to weather, launch delays, and shuttle crew schedules. Nominal duration is usually known in advance, but may increase if there are deployment or checkout problems. These operations have high priority.

Mode 2 - Emergency operations

Both time of occurrence and duration are unpredictable. Emergency operation requires immediate access, for an unknown duration. Several contacts may be required before return to normal operations is possible. Emergency operations usually are given highest priority.

Mode 3 - Target of opportunity

Target of opportunity operations involve observations of unpredictable phenomena. Because these phenomena are often of limited duration, and because it is often important to capture data on the early stages of such an event, science payoff may decrease if access is delayed. Priority will vary, e.g., a supernova is a rare event; while solar flares, although unpredictable, occur frequently.

Mode 4 - Aperiodic Operations

Typical mission operations, consisting of frequent, but not periodic, data collection or commanding sessions. This requires several contacts per day for real time data collection or transfer of recorded data, and for health & safety (H&S) monitoring and commanding. The duration and frequency may vary, due to variations in science objectives, target characteristics, or visibility. Failure to obtain any one contact is of relatively little consequence, but provision of an agreed-to level of service may be required.

Mode 5 - Periodic Operations

An investigation is carried out making the same observations day after day for months or years. Data is transferred from the spacecraft at regular intervals. With occasional missed contacts permissible.

Virtually every mission operates initially in mode 1 for Launch/checkout. When the spacecraft is seen to be operating normally, operation in mode 4 or mode 5 begins. Some spacecraft may occasionally be operated in mode 3 (Target of Opportunity), but few missions plan to operate this way normally, due to present limitations on short lead-time access to the spacecraft via TDRSS. Mode 2 (Emergency) is used only when the spacecraft is in danger. Operations may occasionally revert to mode one to handle orbit maneuvers or on-orbit servicing.

2.4 Service Taxonomy

A variety of characteristics of the service SNC could provide have been identified. These characteristics are not necessarily recommended. Rather they are provided to define the full range of potential services and to provide input to the process of defining the major alternatives for SNC.

SERVICE REQUESTS: Service requests can be made in a variety of ways. Requests may be made for fixed time periods, without specifying any existing flexibility. Alternatively, flexibility in the time at which service is to be provided may be included in the request. Flexibility may be represented as a time window within which all times would be considered equally desirable or as a combination of a preferred time and the periods before and after the preferred time which are acceptable. When time periods are specified they may be described in either relative or absolute terms. Relative time specifications may be expressed as a number of minutes after Acquisition of Signal (AOS), for example.

Further abstractions are also possible in the way in which requests are made, the most abstract form being the generic request. A generic request does not specify hours or dates; rather, it is a statement of the rules governing the selection of hours or dates for those types of service which have some pattern which lends itself to abstraction. Naturally, requests can be made in the more traditional explicit manner, or as some combination of the two approaches in which some elements are specified generically and other elements are specified explicitly.

Alternatives regarding degree of explicitness also exist with respect to selection of spacecraft, antennas, and even type of service [Multi-access (MA) vs. SA]. Service requests can also be independent of one another or be integrated at two levels. At the first level single events which are related to or dependent on one another can be tied together to form a "macro" event. At the second level events associated with multiple spacecraft can be tied together to form a "composite" event. The need for integration of the second type can be found in manned flight when spacecraft / stations need to dock with one another; both spacecraft must have service at the same time, scheduling one without the other is not acceptable. At either level, the integrated request must be treated as a whole, all parts of the request must be scheduled for the request to be satisfied. The last type of service request identified provides the ability to request a switch of service from one TDRSS satellite to another in near real-time. This type of request would provide the ability to maintain nearly continuous service for users with a continuous coverage requirement in the event of a TDRSS related failure.

TYPE OF SCHEDULING: Three alternatives exist with respect to the way in which scheduling is performed. A "static" approach would attempt to produce a schedule at some time in advance and maintain that schedule up to the start of the associated services. A "dynamic" approach would perform little, if any, advance scheduling; requests would be allocated to resources as they arrived. A fluid approach would develop the schedule in advance, but only when necessary and only to the degree required at the time; the schedule would transition from a totally fluid state, through a slushy (partially frozen, partially explicit) state, and finally to a fully frozen condition.

LENGTH OF ADVANCE SCHEDULING PERIOD: In the case of a static or fluid scheduling approach the length of the advance scheduling period could vary from hours to days to weeks. Moreover, the optimal length of this period could be different for different users.

SCHEDULING POLICY: Also in the case of a static or fluid scheduling approach several different policies could be used to allocate resources to requests. Resources could be allocated to the most difficult requests first. Alternatively, the policy being followed today of allocating resources to the highest priority requests could be followed. A variant to both of these alternatives would be to pre-allocate resources to particular users. This would be particularly helpful in the case of STS where the exact TDRSS schedule cannot be known until several hours after launch.

HANDLING OF PRIORITIES: Priorities could be applied in several different ways. One example is to request all parties in conflict to alter or drop their requests in conflict regardless of their priority; priority would only be used as a last resort to resolving conflicts. Such a scheme could be implemented by establishing levels of "service importance" such as normal, critical, and emergency.

HANDLING OF BLOCKED REQUESTS: Blocked requests can be handled in several different ways. First, the request can simply be rejected without any further action. Second, the scheduling system can attempt to find a near fit based on information contained in the request. Third, the request can be placed in a service queue, to be satisfied when resources become available. Finally, coordination with the user payload operations control center (POCC) can be initiated to find an acceptable new time. This coordination can be performed in either a manual voice mode, an automated computer to computer mode, or a combination of the two.

2.5 Functional Analysis

The primary objective of the functional analysis is to establish a baseline set of capabilities for the Space Network and enable the formulation of architectural alternatives. To define a structure useful for this work, the highest level (Level 1) SN functions are categorized as Administrative, Operational, and Engineering functions. The International Standard Organization (ISO) Open Systems Interconnection (OSI) Management Framework (ISO 7498-4) was used as a guideline to decompose the Operational level 1 functions. In particular, these functions relating to management are a one-to-one correspondence to the OSI network management functions.

In the decomposition used in this study, the level 2 functions generally describe the user's view of requirements, i.e., services to be provided by SN. Figure 2.5-1 presents the SN level 1 and 2 functions. The level 2 functions are decomposed further, such that the lowest level functions (level 3 or 4) are simple subfunctions that can be performed by a single computational or organizational entity. The lowest level functions are generally recognizable utility or generic capabilities and document the system analyst's view of requirements, i.e., functions provided by SN. The driving factor in determining the depth of functional decomposition is the need to analyze the alternatives and address the key issues. This decomposition used the applicable ISO/OSI Management Framework to facilitate use of commercial off the shelf (COTS) products by the SN. The complete functional decomposition used in this study is documented in Appendix A.

The key issues to be addressed by this study pertain mostly to Operational functions. Therefore, these functions were analyzed in greater detail compared to the Administrative and Sustaining Engineering functions. The allocation of these functions to various SNC, User Systems, Advanced TDRSS Ground Terminal (ATGT) and other service providers (such as CDOS, NASCOM-ID) and its impact on architectural alternatives are discussed in Section 3.

The SN functions can also be broadly viewed in terms of communications functions and control functions as shown in Figure 2.5-1. The communications functions are those that directly provide the user services, e.g., delivery of bits, while the control functions are those required to support the user services. The communications functions are implemented in the Advanced TDRSS (ATDRSS), ATGT, and ground network (GN) assets while the control functions may be embedded in these communications assets or resident in external systems. The

functions that are not embedded in the communications assets may

I. Administer SN

- A. Co-ordinate Organizational Interfaces
- B. Provide Technical Operations Direction
- C. Manage Human Resources
- D. Perform SN Fiscal Planning

II. Operate SN

- A. Provide User Services
- B. Manage Configuration (including Resource Allocation)*
- C. Manage Faults*
- D. Manage Performance*
- E. Manage Security*
- F. Manage Accounting*

III. Engineer SN

- A. Inter-System Configuration Management
- B. Simulate and Test SN
- C. Maintain SN
- D. Provide SN Training
- E. SN User Liaison

* Open Systems Interconnection Management Functions

Table 2.5-1: Level 1 and Level 2 Functional Decomposition

be further categorized as either intra-SN functions or inter-system functions. This leads to the following functional categorization:

- o subsystem functions - embedded in the communications assets,
- o intra-SNC functions - not embedded in the communications assets and performing only SN control,
- o inter-system control (ISC) functions - functions that involve the control and co-ordination with systems external to the SN,
- o gateway functions - provide the interconnection with the international partners.

In this study, the major focus is on the intra-SN functions that are not necessarily embedded in the communications assets, and they are referred to as the SNC and ISC functions in the following sections. While this is not strictly correct, it simplifies terminology by eliminating the need for an additional acronym.

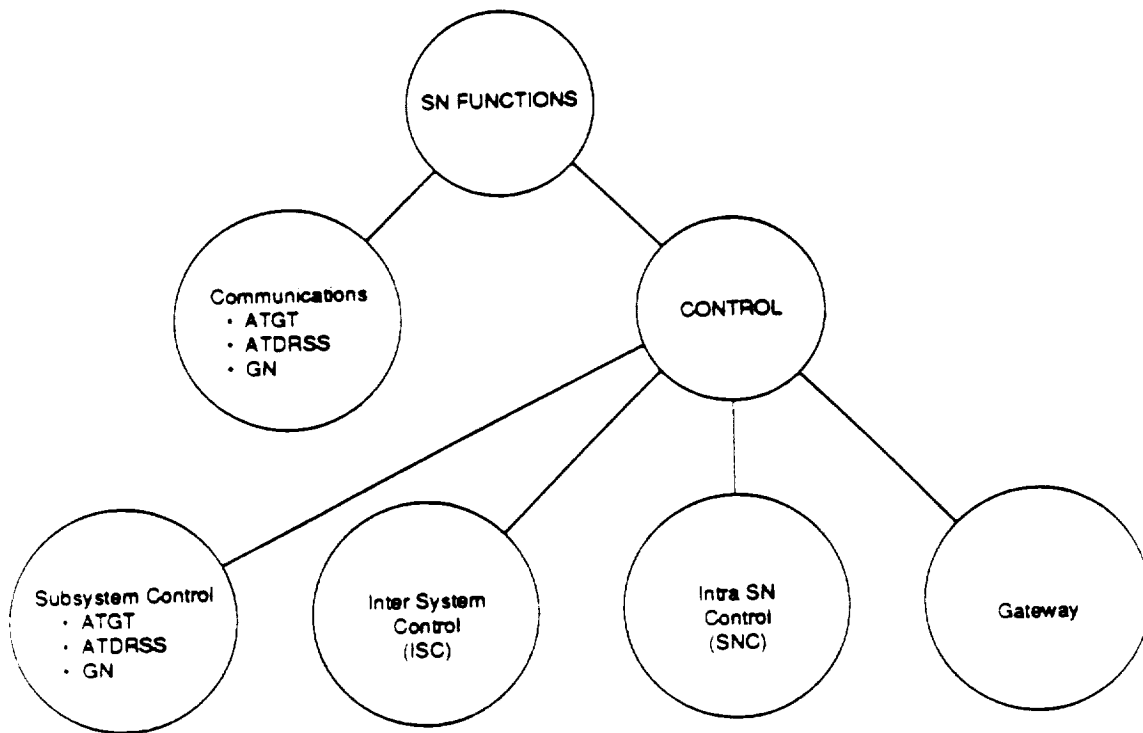


Figure 2.5-1 SN Functional Categorization

3. Architectural Alternatives

In this section a set of architectural alternatives is formulated for Space Network Control (SNC). First, the overall approach to formulating alternatives is presented in Section 3.1. Since resource allocation is the primary area of interest, specific alternatives for this function are presented in Section 3.2. Then a basic set of alternatives is presented in Section 3.3 in terms of:

- o baseline architecture similar to the current system,
- o partitioning the SNC into real-time and non-real-time subsystems
- o partitioning the SNC into classified and unclassified alternatives.

Subalternatives are also formulated in terms of integrating functions with Advanced TDRSS Ground Terminal (ATGT) and performing the inter-system control functions in a stand-alone system or integrated with an external system such as NASCOM II or Customer Data Operations (CDOS). The formulation of these alternatives involved a relative lengthy process and supporting material is provided in Appendix D. This supporting material includes a definition of a control taxonomy, analysis of intra-system control functions, and an analysis of inter-system control issues.

There are number of common functions, referred to as infrastructure, that are needed for support across all of the alternatives. These functions are presented in Section 3.4.

3.1 Approach

As a starting point, the SNC functions are partitioned into resource allocation and system management functions. In the SNC environment, a major driver is the resource allocation, i.e., the algorithms for providing access to the communications assets, so the analysis focuses initially on the resource allocation function and then integrates the other management functions.

The approach for formulating these alternatives is summarized in Figure 3.1-1 and consists of the following elements:

- o system "information interface" for both users and operators
 - manual-automation tradeoffs
 - data input/output techniques
- o resource allocation of communications assets to users - these algorithms may be real-time, preplanned, or hybrid.
- o allocation of functions and data to physical entities
- o system management
 - + fault management,
 - + performance management,
 - + configuration management,
 - + accounting management, and
 - + security management
- o definition of infrastructure requirements, such as communications, to support the allocation of functions and data.

The allocation of functions and data was iterative. It was done for an initial set of SNC resource allocation functions. As the system management functions were defined and allocated, this process was repeated until a satisfactory set of alternatives was formulated.

Specific alternatives were then formulated by taking various options for each of the above elements. This could produce a relatively large number of alternatives, but not all combinations practical or reasonable. Those that are not practical or reasonable will be eliminated from consideration. In formulating these alternatives, it was intended to facilitate the use off-the-shelf components to the extent practical, such as security, communications, and scheduling products.

To identify new technologies and concepts for the SNC alternatives, the capabilities of systems having similar resource allocation requirements were surveyed. This systems included the Department of Defense (DoD) Consolidated Space Operations Center (and other classified sites) and commercial satellite and telecommunications carriers (COMSAT, INTELSAT, GTE SpaceNet, AT&T).

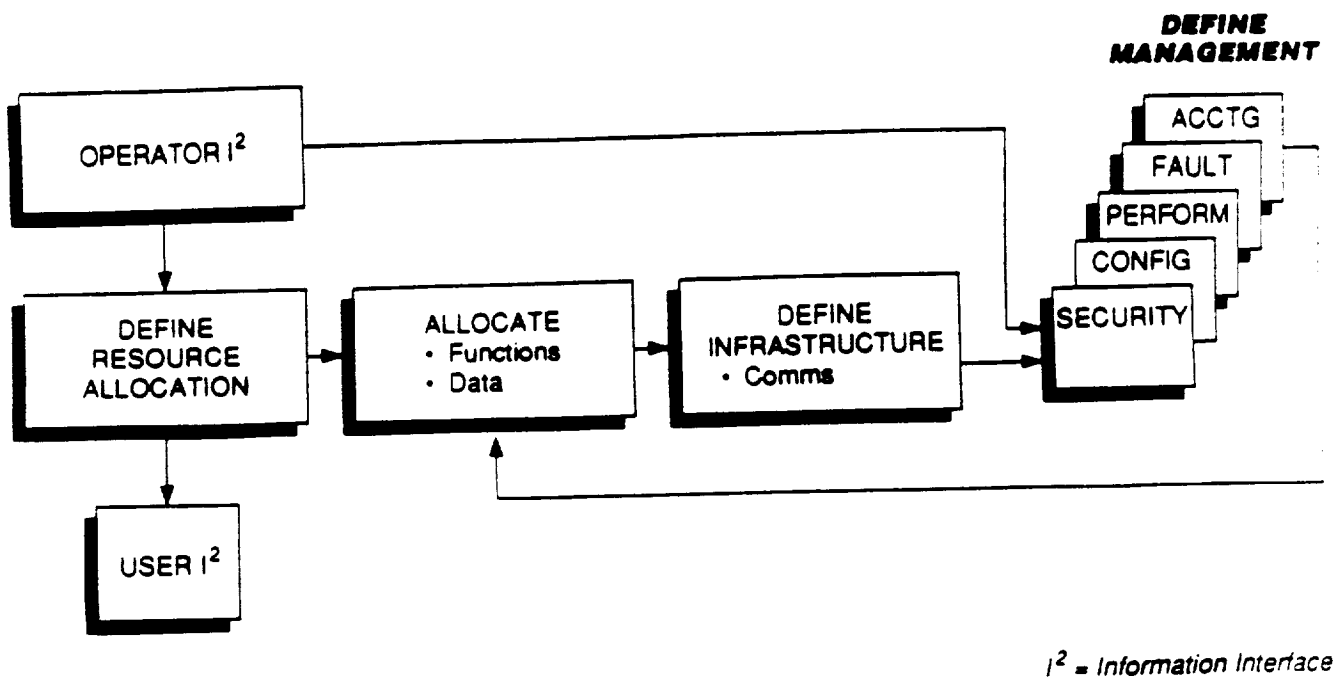


Figure 3.1-1. Formulation of Alternatives

To identify additional systems with analogous requirements, the resource allocation problem was represented in a mathematical abstraction and an electronic literature search was conducted. Based on this abstract formulation and the literature, a number of "abstract analogues" were identified. The results of this survey are presented in Appendix B.

As part of the literature search, recent resource allocation work to determine recent advances in the field were identified. Allocation techniques for job shop scheduling, air traffic control, operating systems scheduling, telephone circuit assignments, mission sequencing, and printed circuit board routing were considered.

3.2 Resource Allocation

3.2.1 Overview

In this section the set of resource allocation algorithms that were considered in this study, depicted in Figure 3.2-1, are presented. As shown in the figure, the key elements of resource allocation are scheduling, access, and resources. To begin, two extreme approaches were formulated for scheduling, pre-planned and demand access algorithms; these alternatives are described in terms of their purpose, functionality, and "information interface" in Sections 3.2.2 and 3.2.3, respectively. The pre-planned approach is analogous to the process currently employed by the Network Control Center (NCC), but incorporates a number of enhancements conceptualized during this study; it consists of two subalternatives, fixed and fluid scheduling. The demand access approach allocates resources as needs occur and is the "telephone network" approach to resource allocation; it consists of two subalternatives, blocked and queued services.

Additional alternatives can be readily defined based on the length of the schedule period. For example, a short term schedule of several hours duration may be kept to enable the resource allocation to be performed more efficiently. This approach is described in Section 3.2.4.

The second element of resource allocation addresses the dedication of resources to user groups or functions. For example, some resources may be dedicated to one user group while the remaining resources are dedicated to another user group. This resource partition establishes subnetworks within SN. Resource partitioning alternatives are discussed in Section 3.2.5.

Then the access element of resource allocation is addressed. A set of hybrid access alternatives based on the characteristics of the users, and traffic types are formulated. In this variant different scheduling techniques may be used within the same resource partition. These alternatives include schemes where:

- o some users may use demand access while others use pre-planned,
- o some traffic types (payload data) may use pre-planned access while other traffic types (spacecraft health and safety) may use demand access),

These alternatives are described in Section 3.2.6.

3.2.2 Pre-Planned

3.2.2.1 Purpose

The purpose of the pre-planned approach is to provide a means by which users of the SN can schedule services in advance and be assured that a subset of the requested services will, in fact, be provided. This approach allows users to plan their requests far enough in advance that the needed access to their satellites can be attained.

As shown in Figure 3.2.2-1, there is a tradeoff between an "abstract scheduling cost" and the time at which resource assignments are frozen. Scheduling cost can be thought of as a function of network utilization and the number of changes resulting from fixing the schedule at some point in advance. If the schedule is fixed very near to the time of service there is insufficient time to resolve conflicts and achieve maximal network utilization. In contrast if the schedule is fixed too far in advance users will find it impossible to faithfully predict their use / needs; an increased number of changes will be required to reflect their ultimate needs. By finding the best tradeoff between these competing factors maximum SN Resources utilization can be achieved with a minimum of effort. The pre-planned approach can also insure that resources are allocated in proportion to the needs / priorities of the individual users.

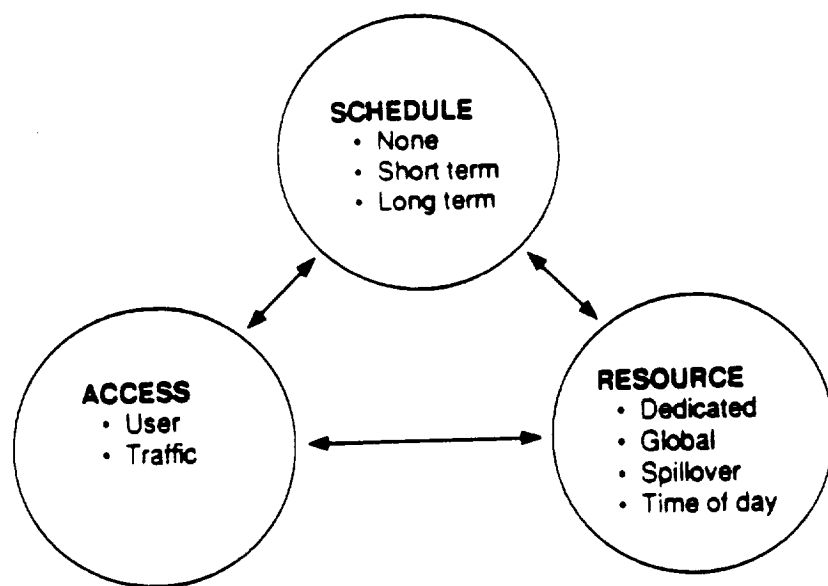
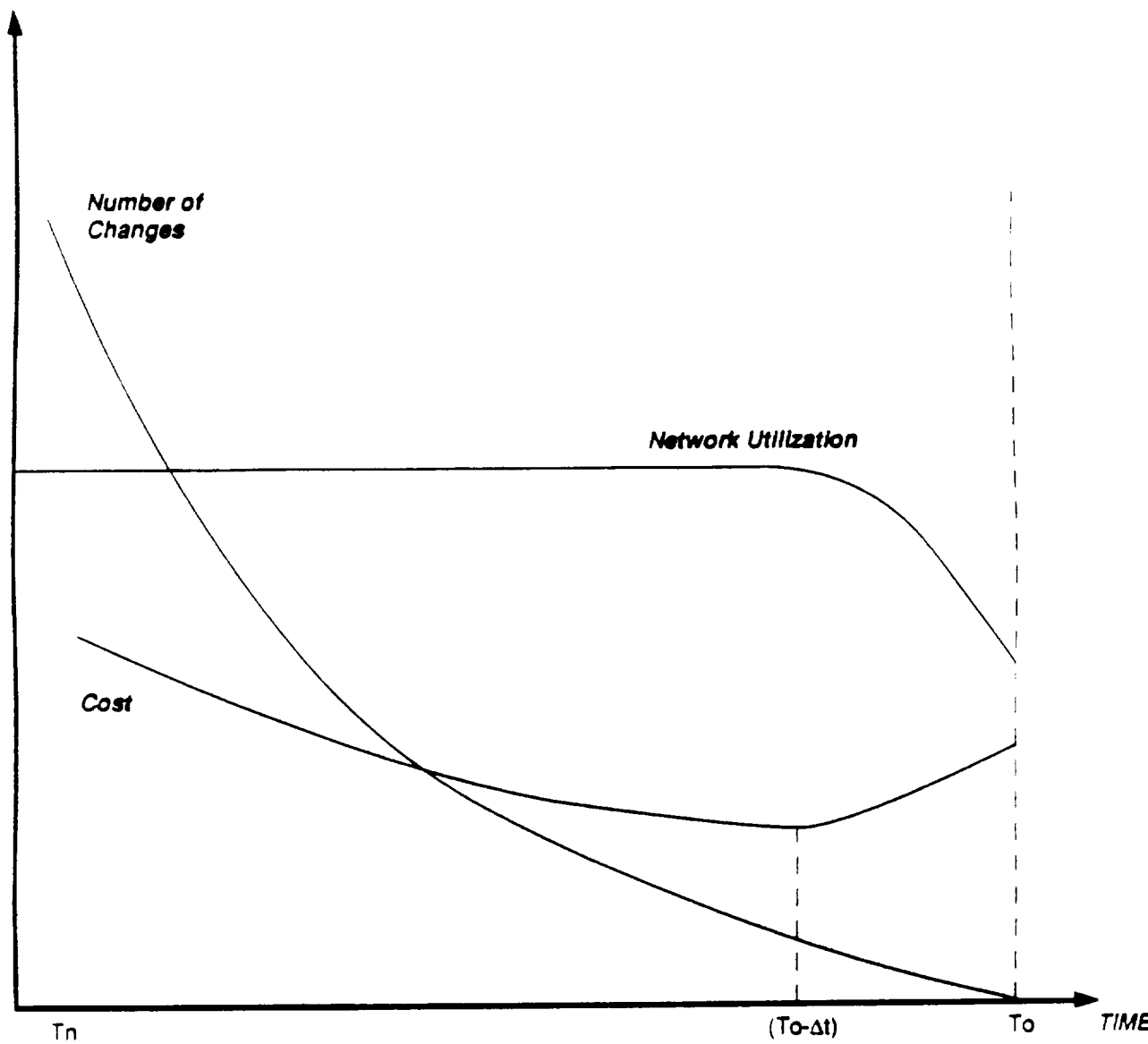


Figure 3.2-1. Universe of Resource Allocation Algorithms



$T_o - \Delta t$ = Optimal Freeze Time
 T_o = Execution Time
 T_n = Start of Schedule

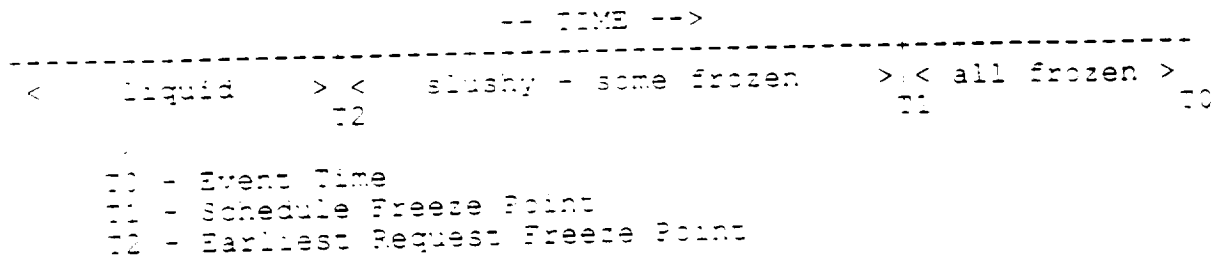
Figure 3.2.2-1. Elements of Scheduling Cost

study; it consists of two subalternatives, fixed and fluid scheduling. The demand access approach allocates resources as needs occur and is the "telephone network" approach to resource allocation; it consists of two subalternatives, blocked and queued services.

Pre-planned scheduling can be performed in two distinctly different ways. The first alternative is to view the schedule as a fluid entity which evolves over time, and becomes more and more fixed as a particular time of service approaches. The second alternative is to view the schedule as an entity which is fixed at a point in time and maintained as an established entity. These alternatives are described below.

3.2.2.2 Fluid Scheduling Functionality

Fluid scheduling can be viewed as a dynamic scheduling approach in which resources are pre-allocated but only to the extent needed (i.e. not prior to the specified freeze point). This parallels the concept of 'just-in-time' scheduling that is now being used in some job-shop environments. Thus, the state of the schedule can be viewed as evolving from liquid, through a partially-frozen (slushy) phase, and eventually becoming totally frozen (at the "schedule freeze point").



Conflict resolution would occur to some extent in every phase; however, the exact meaning of conflict resolution would be different in different phases as would be the strategies for resolving conflicts (see the section describing allocation strategies below). The fluid nature of this approach would be further enhanced by recognizing that Payload Operation Control Centers (POCCs) vary in the amount of lead time needed to prepare for an event and that certain types of requests need more lead time than others. This recognition would be provided by allowing the freeze point to be specified for each request.

3.2.2.2.1 Freeze Point

The primary premise of the fluid scheduling concept is the existence and location of the Freeze Point. The Freeze Point represents the point in time prior to event initiation when resources are reserved or 'frozen' for use by a requester. In fact, in the current NCC operations, the Freeze Point is the time when the planning schedule becomes the active schedule. Once the schedule has been frozen, users can still request changes and additional services, but the resources allocated to scheduled events are not available.

The freeze point for the fluid scheduling concept would work as follows. Initially, a request for service would be received by the SNC. The lead time needed by the POCC to prepare for a scheduled event would be contained in that request. This lead time specifies the "request freeze point". Specifically, if the POCC specifies that it needs a slot on noon Tuesday, but needs 96 hours to prepare for contact, then the request freeze point would be noon the previous Friday. At that time the resources become allocated to the requester with a 'degree of certainty' (most likely a certainty less than 1). However, the allocation could be nullified if a higher priority request was received that had an unresolvable conflict with the initial request. In practice, POCCs would not be allowed to specify any freeze point they desired. Rather, four to six optional freeze points could be selected from. This would accommodate two or three classes of POCCs each having two or three classes of requests. Dividing requests into classes would allow science oriented requests to be differentiated from Health & Safety (H&S) oriented

requests, etc. Differentiating requests by user and type of service may also be related to priority and provide a mechanism for formalization of priority by type of request.

The schedule freeze point then, is the point at which resources are allocated to requests with absolute certainty. After the schedule freeze point, changes and demand access requests can still be processed by the system, but conflicts will be resolved in favor of the 'frozen' request. Manual intervention by the SNC operators would be required to override this default on a case by case basis.

3.2.2.2.2 Allocation Strategy

Another facet of fluid scheduling is the concept of the allocation strategy. The allocation strategy is the method used to assign specific times to user requests. This is in addition to the allocation policy which determines in whose favor conflicts are resolved when two requests can not simultaneously be satisfied. In contrast, the allocation strategy proposes a method of attempting to generate an optimal schedule through the application of one or multiple algorithms. The allocation strategy can significantly impact the User-System Interface, and thus is discussed here.

Three possible allocation strategies have been identified (others may exist as well) and are referred to as:

- o Demand Leveling
- o Window Pruning
- o Constraint Relaxation

The demand leveling strategy would provide a means to smooth out parts of the schedule in which excessive demand existed. The SNC would evaluate all requests submitted and determine, in an approximate manner, where excessive and minimal demand existed. The SNC would then determine which requests could be moved from high demand to low demand parts of the schedule. Demand leveling is based on strategies used by some printed circuit board routing programs to even out the connection density prior to starting the routing operation.

With the Window Pruning strategy, the users would submit requests to the SNC requesting service of a specified duration within a defined window. The SNC would lay out all of these requests and attempt to shift all the requests within their respective windows until all could be satisfied. Unsatisfiable requests would be rejected and resources assigned to the rest. The concern with this allocation strategy is that the number of possible schedules that the SNC is presented with increases exponentially with the number of submitted requests. This could affect the processing power needed to generate efficient schedules in a short (i.e., hours) time frame.

With the Constraint Relaxation strategy, the users submit requests in the same fashion, except they include an optimal time for service with each request. The SNC then attempts to schedule the optimal time for each request. When conflicts between requests are initially detected, the specified time windows are used in an attempt to shift one or both of the requests. In this way, shifts can be done within localized areas, with local optimization replacing global optimization. This should reduce the processing power and time needed to allocate resources. This approach is used by several advance planning systems (e.g. NOAH) for this reason.

In reality, a combination of these three strategies could be used with a fluid schedule to achieve near-optimal allocations with nominal processing power. Specifically, a combination of a demand leveling and window pruning approach would be used during the liquid phase of the schedule (in conjunction with the allocation policy) to reject or shift some percentage of the requests to less populated areas of the schedule. This would represent an attempt to achieve global optimization by spreading out the requests based on a gross level of scheduling granularity. Upon entering the 'slushy' phase of the schedule, a constraint relaxation approach would be used to shift requests within their specified windows in order to assign specific allocation times.

3.2.2.3 Fixed Scheduling Functionality

In the case of fixed scheduling, the schedule would be produced at a preset time. All requests would be submitted in advance of this time or risk not obtaining the desired service. Users would develop their requests independently, without knowing the amount of interest in various time periods. Requests would be submitted in the same form described for fluid scheduling except that the desired freeze point would not be required. Conflicts would be resolved in one pass, to the degree possible, utilizing a constraint relaxation strategy. Shift requests / responses / notices would be used to automate resolution of outstanding conflicts. Potential rejection notices and stand-by requests could be used in the same manner described for fluid scheduling. Further additions / changes to the resulting schedule would be accepted, to the extent that they did not cause a conflict, until shortly before the time of service.

3.2.2.4 Information Interface

The interfaces among the various organizations and elements would be designed and managed to allow the associated systems to interface in a fully automated fashion. These interfaces would replace the voice coordination which occurs today. The information which would need to be exchanged to achieve this goal is described below.

Requests for service made in conjunction with a pre-planned scheduling concept must indicate when service is desired and the type of service desired. An example of such schedule requests is shown in Figure 3.2.2.4-1. Two types of requests would be required. The first type, window / fixed, would support all requests for specific times or time windows. The second type, generic, would support requests from which the SNC would develop specific schedule instantiations.

As shown in Figure 3.2.2.4-1 schedule requests would contain an indication of the optimal time of service to support constraint relaxation. Schedule requests would also contain an indication of desired freeze time. The schedule request would provide a means to indicate any dependence on other pending requests. This would allow, among other things, the ability to indicate that a window request must either precede or follow any generic requests which would be expected to be scheduled in the same time period or the ability to indicate the order that two overlapping window requests must occur in. A single request could contain multiple linked events or event elements and could specify the inter-element dependencies.

An example of the schedule which would be developed by a pre-planned scheduling concept is shown in Figure 3.2.2.4-2. It should be noted that schedule detail records would exist in two varieties, records corresponding to requests still in the planning stages, and records corresponding to requests which have been "frozen" (placed in the active schedule). The schedule would not contain firm times for requests which were still in the planning stages. Similarly, satellite assignments would not be entered for records belonging to requests in the early planning stages. As the request proceeds from stage to stage, the satellite information becomes more specific, starting first as the generic satellite location (i.e. East) or even as a TBD (in which case the scheduling system decides which satellite to use), proceeding to a specific satellite (i.e. East-1), and finally to a specific Single Access (SA) antenna or Multi-access (MA) element.

Move requests and shift requests would contain a suggestion to change the time of service associated with a given request. Move requests and shift requests would be sent in order to resolve schedule conflicts. Move requests would suggest an entirely different time for the service, shift requests would suggest a time only slightly different than the original request. Move requests / shift requests would contain a "response needed by" field, and thus would essentially eliminate what is done manually today.

A move response or shift response would be issued by the POCC system subsequent to its evaluation of each move request and shift request, and could contain three types of information. First, the response could indicate willingness to accept the suggestion. Second, the response could indicate rejection of the suggestion and an indication of the reason for rejection. Third, the response could indicate an alternate or modified suggestion.

Master Request Record

User Satellite - ID	Request Number	Request Type	UDRS Satellite Requirements	Inter Request Dependence	Generic Requirement	Submission Date	Submission Time	Commitment Requirement
3	59	Window	East	None	N/A	90317		3 Days
8	20	Fixed	West	None	N/A	90302		1 Day
8	60	Fixed	None	None	N/A	90317		1 Day
5	2	Generic	None	None	Once Every Orbit	90015		6 Days
5	18	Window	None	Before Generic Request?	N/A	90302		6 Days

Window/Fixed Request Record (used once deleted after event occurs)

Request Number	Element Number	Start Day	Service Type	Service Parameters	Earliest Start Time	Latest Start Time	Duration	Preferred Start Time	Preferred Direction of Shift
59	1	90345	SSAF	_____	16 19 30	16 39 30	10 Min	16 25 00	Forward
59	2	90345	SSAH	_____	Element 1 + 5 min	Element 1 + 10 min	25 Min	Element 1 + 7 Min	Forward
59	3	90345	SSAH	_____	AOS 2ND - Orbit Following	LOS 2nd Orbit Following 20 min	20 Min	Element 1 + 2.5 Hours	Forward
60	1	90341	KSAH	_____	80 40 10	_____	10 Min	N/A	N/A
20	1	90320	MAF	_____	10 15 21	_____	_____	N/A	N/A
18	1	90320	MAF	_____	22 20 10	Per Dependence	9 Min 30 Sec	22 25 15	Backward

Generic Request Record (used repetitively; retained by system)

Request Number	Element Number	Service Type	Duration	Intra Request Dependence
2	1	MAF	10 Min	_____
2	2	MAH	20 Min	Element 1 + 5 Min

Current Day = 90318

Master Schedule Record

User Satellite - ID	Request Number	Request Type	Processing Status	Request Status
3	59	Window	Planning Step 2	Conflict Element 2
8	20	Fixed	Active Schedule	No conflict
8	60	Fixed	Planning Step 3	Conflict Element 1; shift request sent
5	2	Generic	Active Schedule	No conflict
5	18	Window	Active Schedule	No conflict

Schedule Detail Records - Planning Portion

Request Number	Element Number	Start Day	Service Type	Service Parameters	Satellite	Antenna/ Element	Earliest Start Time	Latest Start Time	Duration	Time Uncertainty
59	1	90345	SSAF	—	East	TBD	16:19:30	16:29:30	10 min	±8 min
59	2	90345	SSAR	—	East	TBD	16:24:30	16:39:30	25 min	±8 min
59	3	90345	SSAR	—	East	TBD	19:19:30	19:44:30	20 min	±8 min
60	1	90341	KSAR	—	TBD	TBD	08:40:10	—	10 min	±5 min 20 sec

Active Portion

20	1	90320	MAF	—	W-1	5	10:15:21	—	9 min 30 sec	±15 sec
18	1	90320	MAF	—	W-1	3	22:31:55	—	8 min 20 sec	±10 sec
2-1535	1	90319	MAF	—	W-1	1	22:40:15	—	10 min	±2 sec
2-1525	2	90319	MAR	—	W-1	15	22:45:15	—	20 min	±2 sec

Figure 3.2.4.2. EXAMPLE FULL XIRPLE SCHEDULE

After finding a solution to a conflict within a move / shift response or among several move / shift responses a move notice or shift notice would be sent to inform users of move(s) / shift(s) selected by the scheduling system.

In addition to the / shift requests described above an automated alternative to calling SNC operator to find when requests could be satisfied is needed. This could take several forms. One such alternative is a calendar showing those time periods with available resources and those time periods during which resources are not available (without indicating who is using the resources or why the resources were not available) could be sent from time to time, or as requested to assist users in choosing time windows. Alternatively, a resource status request / resource status response could be utilized to allow POCC systems to determine whether a time period they are considering has available resources of the type needed or not.

Potential rejection notices would be sent to warn users that their request will be rejected unless other users change their requests. Potential rejection notices would be sent when attempts to resolve conflict were not successful. Users could send stand-by requests indicating their willingness to accept the risk of not getting a particular time period, in the hope that schedule changes would make the resource available.

3.2.2.5 Unclassified Schedule

If the schedule were to become unclassified, the way in which scheduling is performed and / or the information interface could be significantly changed. There are three primary alternatives for performing scheduling with an unclassified schedule. These alternatives are described below.

The first alternative is based on the GTE Spacenet booking system. In this concept users would be able to call into the scheduling system and interact with the schedule. Users would be able to see where the conflicts were but not who was responsible for the conflicts. Users would be able to book any open time without any further approvals or interactions.

The second alternative would still require users to submit requests which would be processed and accepted or rejected by the SNC. However, when conflicts occurred, users would have the benefit of knowing which other users were requesting services for the time period of interest and where openings existed in the schedule. Each user would have the option of moving their requests to an open slot and / or of contacting the other users requesting service for the same time period and negotiating for various time slots.

The third alternative carries the second alternative a step further. Information about the time slots being planned and requested by each user would be shared electronically among the users systems. In effect, each user would have the ability to maintain a composite schedule of desired time periods and / or time periods to be requested. Users would have the ability to negotiate with each other electronically. In practice, conflicts would be resolved among the user community before any requests would be submitted. Requests would merely formalize these actions and allow the SNC to allocate resources. In some cases it might be necessary for the SNC to decide who would be given service at a particular time.

3.2.3 Demand Access

3.2.3.1 Purpose

The purpose of the demand access resource allocation approach is to allow users to request system resources at the time they are needed, as with the telephone system. Within this approach there is no concept of a formal schedule; instead, users submit a request for network access once they have made all the preparations needed to contact the user satellite. The goal of this approach is to maximize real-time access to the Space Network (SN) assets without the need to go through an extended and possibly iterative pre-planning cycle.

3.2.3.2 Functional Flow

In order to present the concepts of demand access, this subsection describes the flow of a user request, as depicted in Figure 3.2.3-1, from its generation at the user POCC, through the resource allocation process and finally, through satisfaction or expiration.

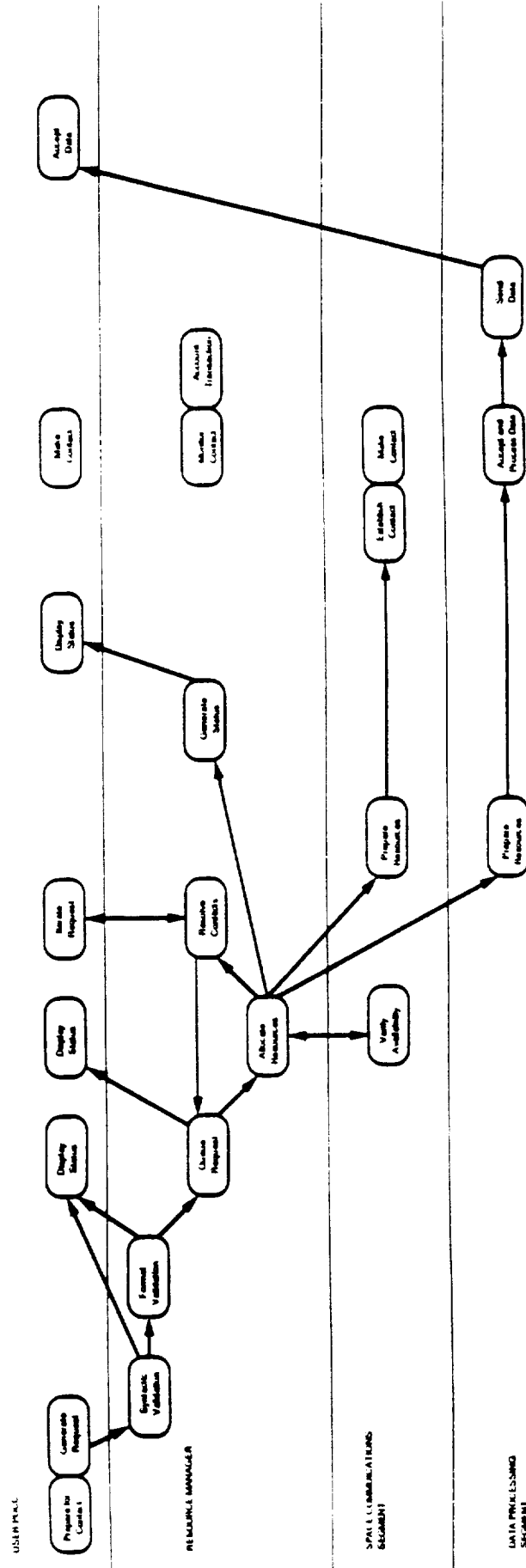


Figure 3.2.3-1: DEMAND ACCESS FUNCTIONAL FLOW

To efficiently utilize a demand access method of acquiring SN resources, the user POCC would first make all preparations for contact with the user satellite. The POCC must assume that access to SN resources can be granted at virtually any time with a low probability of rejection. Specifically, this means that all command loads and other data to be communicated with the user satellite would be produced by the POCC prior to requesting SN access. At the desired time, the user POCC would generate and issue a request for the needed SN resources. The resource request would typically include: the POCC/User Satellites (USAT) id(s), the type of service needed (SA/MA), the direction of service (forward and/or return) the USAT interface parameters (e.g. frequency, encoding parameters, data rate, etc.), the latest possible contact time, and an estimate of contact duration. It is expected that most of this data would be automatically generated and manually verified within the POCC before being transmitted to the SNC.

The request is received by the SNC, where it is automatically validated and processed. Validation is a two step process consisting of an initial syntactic validation (values within legal ranges) and a semantic validation (USAT within view, interface parameters consistent, etc). If the request fails the validation process, it is returned to the user POCC with specific indications of why the request has been returned. The information provided must be sufficient to allow the user POCC to assess the nature of the problem, correct the error and re-submit the request. It is possible, for simple errors, that minimal manual intervention would be required to fix the problem. More serious errors could inhibit request re-submission and require more extensive analysis and replanning on the part of the POCC.

When a request passes the SNC validation phase, it is entered into the queue of pending service requests, and the resource allocator notified that a new request has arrived. The resource allocator scans the queue of requests and determines which one(s) to service. This selection process is primarily driven by the specific resource allocation policy that is in effect at the time. Thus, request may be selected for service based on priority, function to be performed, amount of needed resources, projected duration of contact, or other derived information. The relative importance of each of these factors may vary as a function of time, or be dependent on other events such as shuttle launches.

When the needed resources are available to satisfy the current request, they are immediately allocated to the requester. A notification is then sent to the ATGT and CDOS Segments to inform them of the allocation of the resources. Only when abnormal conditions were detected would the SNC operators need to be notified and become involved in the resource allocation process. Otherwise, normal conditions would cause the various SN segments to prepare to make contact with the specified user satellite. Specifically, the necessary commands would be built and sent to the appropriate Data Relay Satellite to acquire contact with the user satellite. Once initial acquisition, set-up and contact verification has successfully been accomplished, communication between the POCC and the user satellite commences, and commands and data can be transferred for the duration of the contact. Downlinked data will be routed to the Data Processing Segment and/or to the user POCC for processing, as directed in the original request. The Data Processing Segment will independently process the raw downlinked data and transfer the output to the user POCC in an acceptable form upon completion.

Contact termination under a demand access approach is handled the same as with the pre-planned approach. When the contact duration is reached, the POCC notifies the SNC that the connection can be terminated. This information is relayed to the Space Communication Segment so that contact can be relinquished and SN resources can be made available to service other requests. A reply is then sent to the user POCC by the SNC signifying that the request has successfully been completed and providing an account of the level of resource utilization that the request required and the volume of data transferred.

When sufficient resources are unavailable to satisfy a user request, two methods of conflict resolution are possible. First, the request can be considered to be 'blocked' and would be immediately returned to the requester with an indication of this rejection condition. Otherwise, the request can be 'queued' with other validated but pending requests, assuming needed resources will become available shortly. In this case, periodic feedback is automatically provided to the user POCC concerning the status of the pending request. This feedback will include, but is not

limited to, the number of conflicting requests in the queue, and the probability that the request can be satisfied within the specified time window. Note that because it is often dangerous to terminate contact with a user satellite in an uncontrolled fashion, a lower priority request will not automatically be pre-empted to service a newly arrived but higher priority request. However, manual interaction and coordination by the SNC operator(s) may be used to provide this capability in atypical situations.

In addition to status information, it is possible for the SNC to inform the user POCC whether minor modifications to a submitted request would improve its chances for being serviced. For example, the SNC may determine that a request can be serviced if the contact duration is reduced by 5 minutes, or if a MA service is used instead of SA. In this case, the user POCC has the option of leaving the request unaltered, making the suggested changes, or withdrawing the request.

Due to peaks in network demand, it is possible that some requests may not be serviced within the specified time frame. Eventually, the request will expire (pass the specified connect time window) and will be returned to the user POCC unserved. The user POCC can assess this situation, take appropriate action, and issue an updated request. In the worst case, this means that the POCC will need to regenerate command loads for the next feasible contact with the user spacecraft (hours to days later for complex, time sensitive command loads).

3.2.3.3 Information Interface

The information interface required to support a demand access resource allocation approach is simpler than that needed to support the preplanned approach. The difference is that demand access does not require an initial time of contact to be provided in any form (it is assumed to be the time of request submission). The latest contact time and expected duration of the request is still needed to support the queueing of initially blocked requests.

Shift requests and related notices are also used by the demand access approach. However, since queued requests are not supported in the scheduled approach, additional interaction between the SNC and the user POCC is required to report their status.

3.2.4 Short-Term Scheduling

3.2.4.1 Purpose

The purpose of the short-term scheduling resource allocation approach is to provide quick (but not immediate) access to SN resources in an attempt to simultaneously maximize satisfaction of user requests and minimize the number of schedule changes that are needed. The short-term scheduling approach is based on the allocation of resources in a real-time fashion through the periodic creation and distribution of a short term (2 to 6 hour) schedule. The goal of this approach is to provide the users with a familiar, but more timely mode of operation for the allocation of SN assets.

3.2.4.2 Functional Flow

This subsection discusses the flow of a user request from its generation at the user POCC through the scheduling process.

The short-term scheduling approach provides the user a capability between a pre-planned and demand access system. Specifically, it allows the user POCC to request SN resources a short time (approximately 2 to 6 hours) prior to the time of need. This reduces the overall planning cycle, but provides limited time to prepare for contact with the user satellite.

Under the short-term scheduling approach, the user POCC will create a request for service and send it to the SNC as discussed in Section 3.2.2, Pre-planned. When the request arrives at the SNC, it is validated and placed in the queue of requests to be scheduled. The requests are then prioritized, based on the scheduling policy, the specified event start time, and resource availability. The highest priority entry is removed from the queue and merged into the evolving schedule.

When the user request is successfully scheduled, the requester is so notified. The notification will include the start time of the scheduled event, its duration, and other related

information. The user POCC can then complete preparations for contact with the user satellite. Because of the short time frame, it is likely that portions of the preparations may have already been done. At the scheduled time, communications between the user POCC and satellite will be established. The duration of the contact will be as specified in the original event request.

When a user request can not be satisfied because of insufficient resources, the user is so notified. In this case, the SNC will include with the returned status, a set of possible time shifts that would increase the probability of successful event scheduling. As previously discussed, it is then up to the user POCC to assess, modify and re-submit the request to the SNC. If the request can not be shifted, then attempts will be made to shift (or reschedule) the blocking events when they are of lower priority. As much of this process as possible will be automated, only involving the SNC operators and the POCC users when automated methods fail to resolve the conflict. This processing and interaction may continue over several cycles.

3.2.4.3 Information Interface

The information interface needed by the short-term scheduling approach is nearly identical with that specified for the pre-planned approach in Section 3.2.2 of this report. The only difference is that the time between submission of a request and the contact time is shorter.

3.2.5 Resource Partitioning

3.2.5.1 Purpose

The SN will provide service to a diverse set of users having widely different requirements. The purpose of resource partitioning is to isolate the effects of these diverse users.

3.2.5.2 Functionality

In the Resource Partitioning techniques, Advanced Tracking and Delay Relay Satellite System (ATDRSS) assets are dedicated for specific purposes. Within a partition any of the above scheduling techniques could be utilized, and different scheduling techniques could be utilized for each partition.

There are a broad range of resource partitioning alternatives to consider. For example, the MA and SA resources could be partitioned among user groups. In this partition two user group could be identified, A and B, and assigned a set of MA and SA resources. Thus, when attempting to obtain ATDRSS support, the two groups would contend for resources only among themselves and not compete with each other. One way to identify user groups is by activity, e.g., manned and unmanned space flight. In this partition, dedicated Single Access Forward (SAF) and Single Access Return (SAR) channels may be dedicated to shuttle on an east Tracking and Data Relay System (TDRSS) and west TDRSS.

There are a number of variants to this technique to consider. First, the partitions may be time variant analogous to the time of day routing performed by the telephone network. For example, prime time and non-prime time partitions may be defined.

Second, the partitions could share resource on an "as available" basis. In this case, pre-planned access would be used for scheduling both partitions such that no user outside of the designated group would be allowed to schedule support on the dedicated assets. However, demand access for users outside of the user group may be allowed; this is referred to as "spillover". For example, if a user outside the user group submits a demand access request but is blocked on his own assets, service (if available) may be provided on the assets of another user group. Resource partitioning is typically thought of in terms of scheduling, but it also applies to a proven demand access approach. However, in this case no "spillover" would be allowed, otherwise the resources would not be partitioned.

The ramifications of this alternative are very broad and affect the ground systems architecture. For example, the implementation of the resource allocation function may also be partitioned by establishing separate SNC subsystems corresponding to the partitions. These resource allocation subsystems could operate as independent subsystems. However, if "spillover" were permitted or if they were serving as backup for each other in case of failure,

they would have to communicate and would not be independent. Various SNC alternatives based on this concept are discussed in Section 3.6

3.2.5.3 Information Interface

When resources are partitioned, each set of resources will have its own scheduling algorithm. The functional flow and information interface will be very similar to that as if the technique were being applied globally across all resources. Thus, these issues are not discussed in this section.

3.2.6 Hybrid Access

3.2.6.1 Purpose

The purpose of the hybrid access resource allocation approach is to provide the capability to simultaneously support a combination of scheduled and demand access requests. This is similar to the air traffic control environment where small aircraft must be granted a landing slot outside of the scheduled runway usage by commercial flights. This approach is predicated on the belief that there is a dichotomy of request types generated by the user POCCs. Specifically, some subset of requests need extended planning periods with a known interval of contact, while others do not. Hence, the goal of this approach is to support both types of requests in a manner that provides for optimal utilization of the SN assets.

3.2.6.2 Functional Flow

This subsection discusses the flow of user requests from its generation at the user POCC, through the scheduling and/or resource allocation process and finally, through satisfaction or expiration.

When the user POCC prepares a request for service, a decision must be made concerning whether this request is to be scheduled or handled via demand access. If the requested contact will take an extended period of time to prepare for, or is time sensitive, then the user POCC would construct a schedule request. Examples of functions that need to be scheduled include the building of extensive command uploads for complex activities or preparing to accept real-time data from a satellite that has no tape recorder. In these cases, the schedule request need only specify: the POCC/USAT id(s), the type(s) of service needed (SA/MA), and the acceptable constraints on the time of contact with the user satellite. The USAT interface parameters are not needed until just before actual contact with the satellite.

If the requested contact is not time sensitive, then a demand access request can be constructed by the user POCC. The demand access request would be used to make contact with the user satellite for health and safety reasons, or for other short term contacts. In this case, the request must specify all the needed satellite contact information including interface parameters (e.g. frequency, encoding parameters, data rate, etc.).

The request is received by the SNC, where it is validated and processed, as discussed in Section 3.2.3. Once validated, a request to be scheduled is placed in the queue to be scheduled. Because of the mixture of schedulable and demand access capability, the schedule maintained by the SNC defines two distinct periods: the scheduling period and the active period. The scheduling period is that period of time where incoming requests are integrated into the evolving SNC schedule with some degree of certainty. At the boundary between the scheduling and active periods, the schedule becomes 'frozen' and all scheduled requests are finalized. Demand access requests are accepted to fill in any unallocated spaces in the active schedule.

The processing of demand access requests is similar to that defined in Section 3.2.3, Demand Access. The difference is related to the fact that scheduled requests will have already been assigned resources within specified time slots. The resource allocator that processes demand access requests must therefore fit them into the remaining open slots. Movement of the scheduled requests to accommodate a demand access request would be limited and might only be accomplished through manual intervention. Once a demand access request is successfully scheduled, a confirmation message is returned to the user POCC acknowledging the status of the

request. The handling of unschedulable requests or expired demand access requests is the same discussed in Sections 3.2.2 and 3.2.3.

Prior to the start of a scheduled event, the user POCC would be queried by the Space Communications Segment (or the SNC on its behalf) to provide the necessary interface parameters. This action serves to confirm the communication path with the user POCC as well as limit the information that the SNC must maintain for each scheduled request.

3.2.6.3 Information Interface

The information interface for the hybrid access allocation approach is a combination of the scheduled and demand access interfaces specified in Sections 3.2.2 and 3.2.3 respectively.

3.3 Architectural Alternatives

3.3.1 Overview

This section formulates a set of architectural alternatives for the SNC and Inter-system Control (ISC) system(s). As the foundation for this formulation, Appendix D documents a control taxonomy and functional allocation analysis for SN. As defined in the appendix, the six aspects of the control taxonomy are:

- o Functions of control,
- o Elements of control,
- o Data Processing modes,
- o Decision-making modes,
- o Redundancy of control entities, and
- o Location of control entities.

This study used the ISO/Open Systems Interconnection (OSI) framework for "functions of control" for the level 2 functional decomposition of SN operational control requirements, as documented in Appendix A. The lower level functions (level 3 and below) are further decomposed in terms of the "elements of control" - monitoring, data processing and decision making.

The functions of control have been defined in terms of the OSI Management functions and then allocated to the four control entities described earlier in Section 2.5:

- o subsystem control,
- o intra-SN control,
- o inter-system control
- o gateway.

Table 3.3-1 provides a high-level summary of the level 2 and level 3 functional allocation, with the SNC functionality discussed further in Section 3.3.2.

This section synthesizes the information in Appendix D, to formulate architectural alternatives and addresses the redundancy and location aspects of the taxonomy. The set of proposal architectural alternatives is shown in Figure 3.3-1. The primary architectural factors that distinguish these alternatives are the different partitioning approaches for the system functions and databases. The three alternatives are:

- o Integrated
- o Real-time vs non real-time partitions
- o Classified vs unclassified partitions.

Variants of these alternatives can be defined based on the level of integration into external systems. The specific variants considered are integration of the ISC; and the real-time intra-SNC functions. Other factors of the architectural alternatives, addressed in describing the alternatives are the location and redundancy of the components.

Although a large number of candidate architectures can be formulated using different combinations of these architectural elements, four representative architectural alternatives have been selected that provide a basis for discussion of trade-offs in Section 4. These alternatives are not intended to be an exhaustive list of possible and practical architectures, but all four are capable of performing all SNC/ISC functions. These architectural alternatives are discussed in Sections 3.3.3 through 3.3.7.

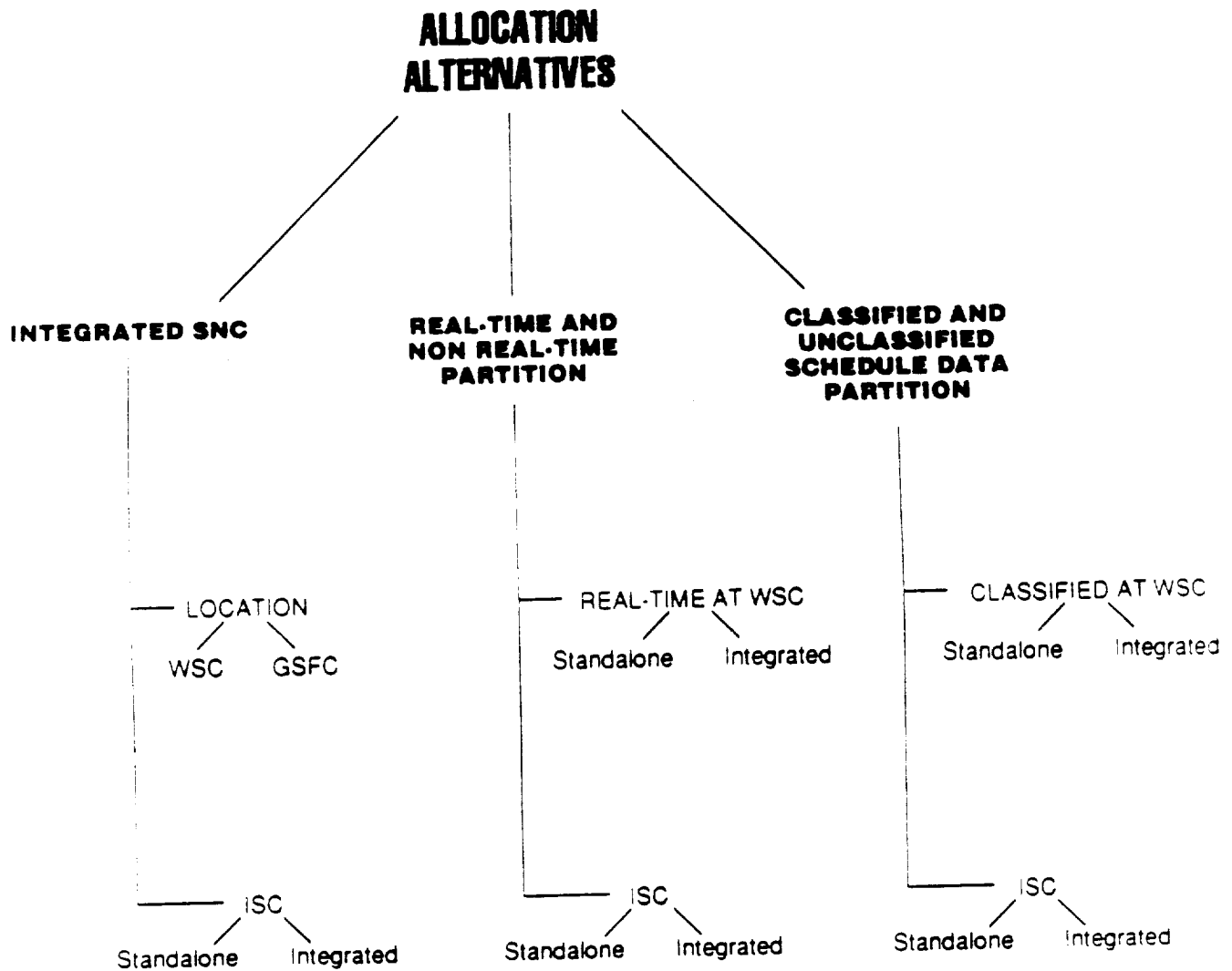


Figure 3.3-1 Summary Resource Allocation Alternatives

3.3.2 SNC/ISC Functionality

As discussed in Appendix D, some control functions are more effective if performed by the subsystems, while others must be performed by a central control entity (SNC for the SN system and ISC for the "System of Systems"). Centralization of control functions is desirable for:

- o functions dealing with end-to-end management, and
- o functions that require arbitration among peers.

SNC/ISC Functions

B. Manage Configuration

1. Maintain SN resource allocation Rules Database
2. Maintain Planned Resource Availability Database
3. Maintain Preplanned Service Request and Scheduled Service Event Databases
4. Manage On-demand Service Request

C. Manage Faults

3. *Manage Inter-system Faults*

D. Manage Performance

3. *Monitor end-to-end real-time performance*

E. Manage Security

1. Manage SN Security
3. *Manage inter-system security*

F. Manage Accounting

1. Maintain SN Rate Database
2. Maintain and report SN resource utilization
3. *Report end-to-end Resource Utilization*

Italics denote inter-system control and co-ordination functions

Table 3.3-1

For the National Aeronautics and Space Administration (NASA) "System of systems", centralization of control functions occurs at multiple levels leading to a hierarchical control system. At the highest level, the ISC system provides control/coordination among the various autonomous systems (such as the SN, CDOS and NASCOM). The ISC system depends on the system managers (such as the SNC for SN) to provide the necessary management information to support the system specific control functions. Thus the system managers perform the dual roles of a manager of their own systems and agents for the ISC system.

The hierarchy of control continues within each system. For example, the SN consists of (at least) six autonomous subsystems that provide user services - two Advanced TDRSS Ground Terminals, three Ground Network (GN) subsystems and the gateway for international partners. Each SN subsystem will be controlled by its own subsystem manager [e.g. ATGT TDRS Operations Control Center (TOCCs)], which will also act as the agent for the SNC system. The focus of this study is the ISC and the intra-system control for the SN only, i.e., SNC. These control functions are summarized in Table 3.3-1.

3.3.2.1 Peer-to-peer Management Approach

The peer-to-peer approach, applied to the NASA "System of Systems", will require the system managers for the autonomous systems to work cooperatively with other system managers for services across system boundaries. With this approach, there is no need for a separate ISC system. All data collection (monitoring) is performed by the individual system managers. The peer managers interact with each other as equals for all decision making and all decisions require agreement (concurrence) between (among) two (or more) peers. If agreement is not reached, peer-to-peer activities cannot be performed, and may result in lack of acceptable end-to-end services (i.e. services requiring use of multiple systems). The primary defect of this approach is that no one system has the complete picture of the end-to-end performance of the network. This significantly complicates fault isolation.

Second, the lack of deterministic procedures to assure end-to-end service is unacceptable for the NASA "System of Systems". The net result of such an approach is that the User System Managers in effect become the de facto inter-system coordinators for their own service events, since ultimately they are responsible for the success of their project. This may lead to each user system implementing its own ad-hoc inter-system control procedures and mechanisms at a significantly higher cost.

Peer-to-peer management approach is very effective for the operation of OSI layers 1-3. The layer operation protocols include well defined control information and procedures to resolve control interactions. Some examples of control information in layer protocols are:

- o packet sequence IDs (fault and accounting management),
- o error detection Cyclic Redundancy Cycle (CRC)/correction codes (fault management),
- o flow control information (performance management), and
- o time stamps (performance management).

At this time (and the foreseeable future), deterministic peer-to-peer layer management protocols do not exist for higher layers. In fact the ISO/OSI standard organization favors system management approach over layer management approach. ISO 7498-4, section 6.2.2 states - "(N)-layer management protocols should only be used where special requirements dictate that system management protocols are inappropriate or when systems management protocols are not available. ... This standard does not require the development of (N)-layer management protocols for each of the seven layers."

3.3.2.2 Managed Objects and MIBs

The control functions listed in Table 3.3-1 were defined within the ISO/OSI Management Framework, as specified in ISO/IEC 7498-4. The interface between a manager and its agents uses the encapsulation principle. The agents maintain Management Information Bases (MIBs), which are a set of well defined managed objects. The agent functions are then defined as alterations or inspections of the "Managed Objects". Imperative command cannot be issued to an agent.

The "Managed Object" approach allows the monitoring of a system state at any level of detail (depending on the MIB definition) by polling for appropriate information. A limited number of unsolicited messages (traps) control the timing and focus of the polling. The goal is interface simplicity and minimization of the management traffic.

The alteration functions can be used by a Manager to control the system state to any level of detail. The exclusion of imperative commands is not as limiting as it may seem at first, since an imperative command can be realized by setting a parameter value that triggers an action. An extreme example would be setting a parameter indicating the number of seconds until system reboot. The MIB concept allows for standard MIBs [defined by various national and international standard organizations, such as ISO, Consultative Committee for International Telegraph & Telephone (CCITT), Consultative Committee for Space Data Systems (CCSDS)], as well as, experimental and private (enterprise) MIBs.

Figure 3.3-2 provides a logical view of the relationship between the ISC and system managers. As shown in the figure and discussed in Appendix D, some control functions can be

potentially realized by using peer-to-peer layer management approach, and does not use the system managers and agents. Instead, it is accomplished by direct exchange of control information between peer layer entities.

3.3.2.3 Example MIB

An example MIB is shown in Table 3.3-2. This was extracted from the Internet RFC 1158, Section 5.2, the Interface Group. A MIB similar to this may be defined for the interfaces between various autonomous systems. Examples include interfaces between:

- o NASCOM II Gateway and a User System.
- o NASCOM II Gateway and CDOS/Data Delivery Service (DDS),
- o NASCOM II Gateway and SN/Data Interface System (DIS),
- o SN/DIS and CDOS/SN interface function (SNIF)

The objects in the example MIB illustrate the concept of reporting by exception. For example, the ifInDiscards object is used to collect information pertaining to buffer bottlenecks. The manager can set a trap, such that the agent will send a message whenever the ifInDiscards exceeds N.

The objects in the example MIB also illustrate the concept of end-to-end fault management, that cannot be performed by individual system managers. For example, the ISC Manager can compare the ifInOctets and ifOutOctets reported by two systems communicating with each other (such as CDOS/Communication Interface Functions (CIF) and NASCOM Gateway). A mismatch of two numbers would indicate a possible fault event. The ISC Manager can then advise the affected ISC agents, query for additional information and initiate diagnostics actions to detect, isolate, and correct the fault.

3.3.3 Architecture 1 - Current approach

This architecture (Figure 3.3.-3) is similar to the Second TDRSS Ground Terminal (STGT) architecture and was selected for this reason. Today, the SNC functions are partitioned between the NCC at Goddard Space Flight Center (GSFC) and the NASA Ground Terminal (NGT) at White Sands Complex (WSC). The NCC is responsible for scheduling SN (TDRSS and GN), Sensor Data Processing Facility (SDPF) and NASCOM services, i.e. end-to-end services. The inter-system control functions are mostly performed manually, while the intra-SN control functions are generally automated. In the STGT era, the ground terminal will be a highly redundant and automated system responsible for performing intra-SN control functions only, e.g., an SN end-to-end connectivity test at the beginning of each service event.

Architecture 1 centralizes all SNC and ISC functions in a redundant classified system in one location (possibly one building) and is independent of the system location. It can be located at WSC or GSFC. The impact of location is primarily in the area of communication costs (see Appendix F) and facilities costs. If the selected location were to shut down due to a catastrophic event (a severe snowstorm, fire, earthquake etc.), the control system could not continue to operate and the ATGT would provide limited functionality back-up. This can be a significant limitation, since the current STGT design does not to provide for back-up operation over an extended period (several days/weeks/months) of time.

These messages and mechanisms are generally used for fault and performance monitoring. The tracer messages can originate either at the user spacecraft (for future spacecrafts where such capability can be implemented) or at the ATGT for the space-ground return links, with the user system as destination address. Similarly, for the ground-space forward links, these messages can originate at the User System NASCOM II gateways with either the ATGT or the user spacecraft as the destination address. The tracer messages are time stamped by the end systems, as well as intermediate systems (i.e. CDOS and NASCOM gateways) and generally collected by the originating system.

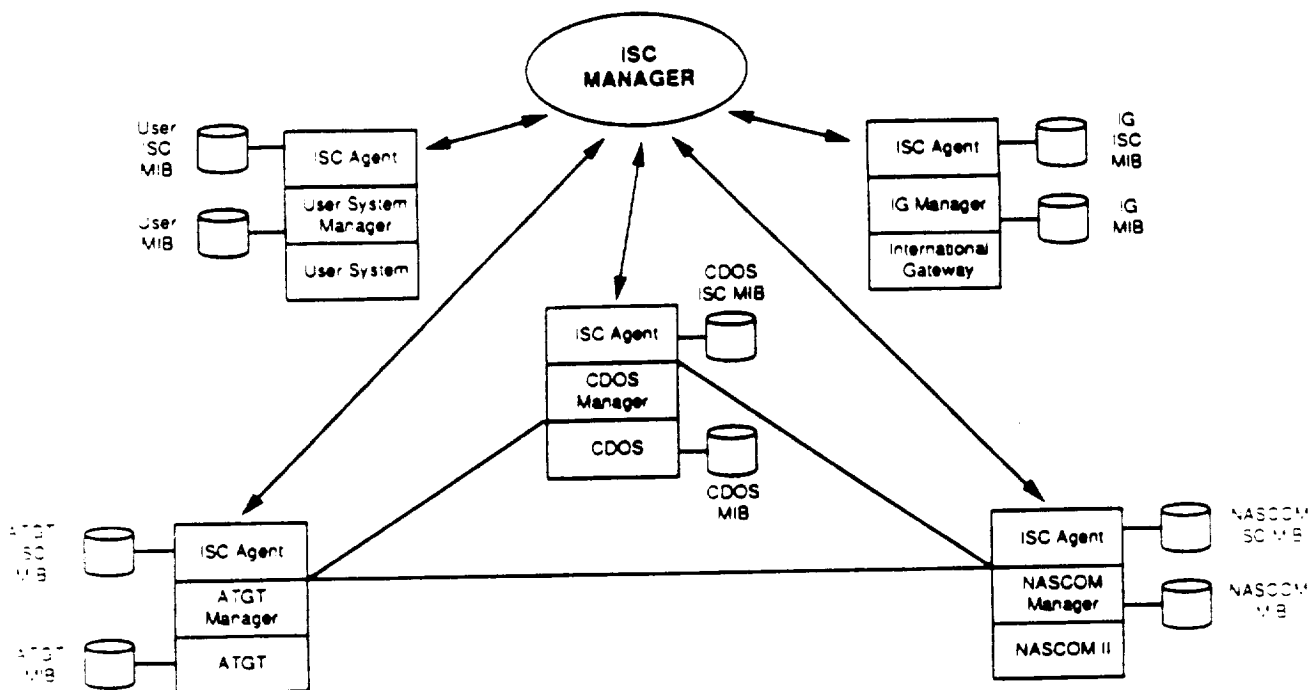


Figure 3.3-2: Inter-System Control — Logical View

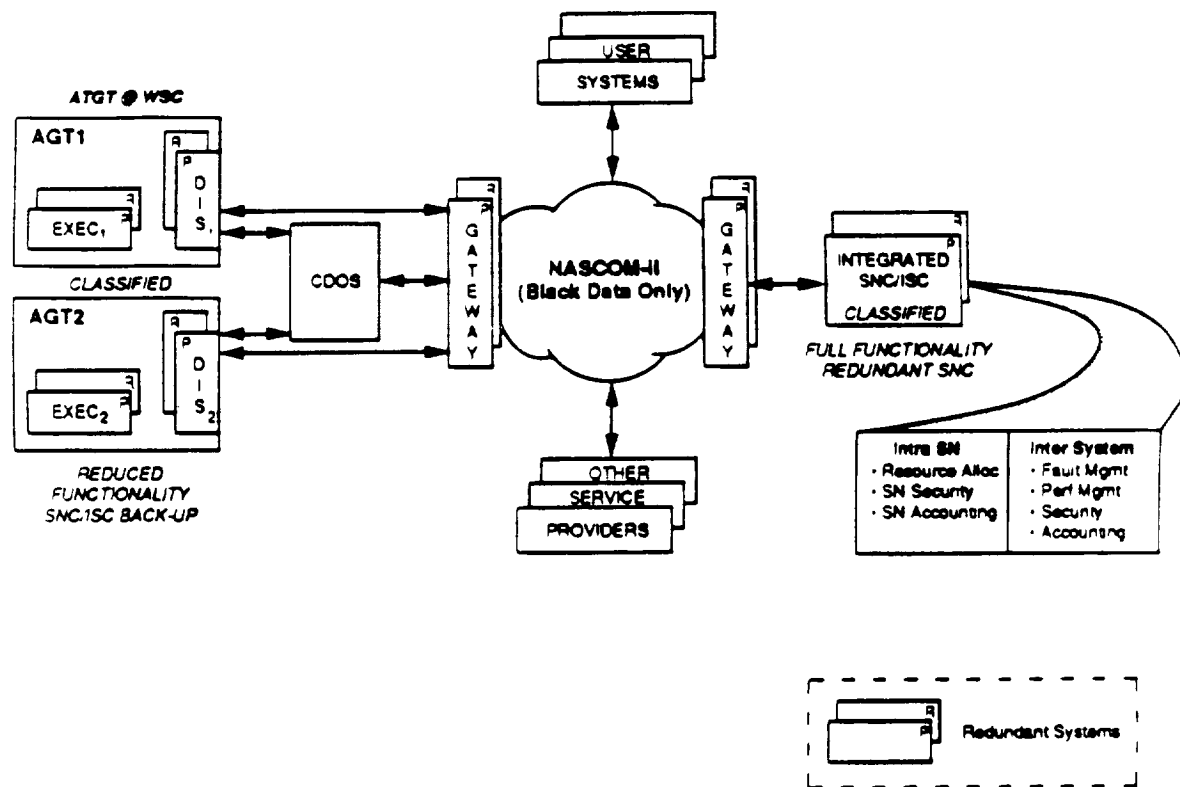


Figure 3.3-3: Architecture 1- Current Approach

OBJECT Name	Object Definition
ifNumber	The number of network interfaces (regardless of their current state) on which this system can send/receive Internet Protocol (IP) datagrams.
ifTable	A list of interface entries. The number of entries is given by the value of ifNumber.
ifEntry	An interface entry containing objects at the subnetwork layer and below for a particular interface.
ifType	The type of interface, distinguished according to the physical/link/network protocol(s) immediately "below" IP in the protocol stack.
ifMTU	The size of the largest IP datagram which can be sent/received on the interface, specified in octets.
ifSpeed	An estimate of the interface's current bandwidth in bits per second. For interfaces which do not vary in bandwidth or for those where no accurate estimation can be made, this object should contain the nominal bandwidth.
ifPhysAddress	The interface's address at the protocol layer immediately "below" IP in the protocol stack.
ifAdminStatus	The desired state of the interface (up, down or testing, i.e., no operational packets can be passed).
ifOperStatus	The current operational state of the interface.
ifLastChange	The value of sysUpTime at the time the interface entered its current operational state.
ifInOctets	The total number of octets received on the interface, including framing characters.
ifInUcastPkts	The number of (subnet) unicast packets delivered to a higher-level protocol.
ifInNUcastPkts	The number of non-unicast packets delivered to a higher-layer protocol.
ifInDiscards	The number of inbound packets which were chosen to be discarded even though no errors had been detected to prevent their delivery to higher-layer protocol. For example, packets discarded to free up buffer space.

Note: The example does not show the object syntax, access and status information.

Table 3.3-2 Example Interface MIB

The originating system computes the transit delays and monitors the running average (over a certain period). If it detects excessive transit delays [compared with agreed to Quality of Service (QOS) thresholds], it sends a request for corrective actions to the ISC system, which analyzes the information (and possibly collects additional measurements) to determine the offending system. It then sends a request for corrective action to the offending system and advises other systems as a precautionary measure. Often, the offending system has its own performance/fault monitoring mechanisms, and it may have already detected the problem.

3.3.4 Real-Time vs Non Real-Time Functions

The functions listed in Table 3.3-1 can be partitioned according to their time criticality. Table 3.3-3 shows the real-time and non real-time partitions for the SNC/ISC functions. Some functions, such as "Maintain SN resource allocation Rules Database (function II.B.1) are not time critical, i.e. do not require real-time or near real-time processing. Other functions, such as "Monitor end-to-end real-time performance (function II.D.3)" are time critical. The data collection and processing must be performed in real-time, while the decision-making and direction (corrective actions) must be performed in near real-time.

The "Monitor end-to-end real-time performance" function is an example of a real-time function that provides continuous monitoring and exception event reporting. It monitors the transit delay using "tracer messages" throughout the duration of each service event. The term "tracer message" is used as a conceptual generalization of a variety of protocol specific mechanisms for measuring transit delays. Some example are:

- o TCP/IP PING (Packet Internet Groper),
- o TCP/IP timestamp request message, and
- o X.400 Probes.

Real-time systems tend to be complex and expensive to develop, deploy, and operate. In addition, the real-time control functions require higher availability, generally achieved by a redundant architecture. This study postulates that the real-time and non real-time functions will be implemented on separate systems communicating with each other over a Local Area Network (LAN) (if the two systems are collocated) or over a Wide Area Network (WAN) (if the two systems are not collocated). Implementing the SNC/ISC system in this manner will lower development, deployment and operating costs.

The architectural issue to be addressed is collocation (or not) of the real-time and non real-time control systems. Architecture 1 (current approach) collocates the two systems at GSFC. Architecture 2 (Figure 3.3-4) locates the real-time SNC/ISC system at WSC and a non redundant non real-time control system at GSFC. Appendix F provides a dataflow analysis for time critical and non time critical messages sent or received by the SNC/ISC systems. It shows that communication costs will be somewhat lower if the real-time control system is located at WSC.

Architecture 2 centralizes all real-time SNC/ISC functions in a redundant classified system at WSC. All non real-time SNC/ISC functions are centralized in a non redundant system at GSFC. Thus it is similar to the current architecture, in terms of lack of location redundancy with one significant difference. The real-time SNC/ISC system is at WSC and the redundancy of this system will be as good as the redundancy of the ATGT subsystem. If the ATGT complex is shut down due to a catastrophic event, there is very little left to control in real-time.

The STGT can provide an existing robust real-time platform for implementing the real-time control functions. Therefore, the real-time control system can be a standalone system (Figure 3.3-4) or it can be integrated with ATGT (i.e., upgrade the STGT to an integrated ATGT by adding the real-time SNC and/or ISC functions to it). The pros and cons of the later approach are discussed in Section 3.3.7.

A non-redundant system is postulated for the non real-time system at GSFC assuming that the system can be repaired in a short (1-2 hours) time period. During this period, preplanned service scheduling will be suspended. If this turns out to be unacceptable, the non real-time system can be made redundant to increase availability.

3.3.5 Classified vs Unclassified Subsystems

The partitioned architecture is based on the premise that it is possible to partition the SN resources and services such that the schedule for a subset of SN resources can be unclassified. The rationale for suggesting this partitioning includes reduced schedule negotiations between SNC and users and lower costs. The control system for the classified subnet can be located within the ATGT classified facility while the control system(s) for unclassified systems can be located anywhere.

Architecture 3 (Figure 3.3-5) centralizes all SNC and ISC functions in two separate, redundant control systems - SNCA at WSC and SNCB at GSFC. The SNCA system is located within the classified facilities at WSC. The SNCB will be a unclassified system that can be located anywhere.

Real-time SNC/ISC Functions

- B. Manage Configuration
 - 4. Manage On-demand Service Requests
- C. Manage Faults
 - 4. *Manage Inter-system Faults*
- D. Manage Performance
 - 3. *Monitor end-to-end real-time performance (except long-term trend analysis, planning input)*

Non real-time SNC/ISC Functions

- B. Manage Configuration
 - 1. Maintain Rules Database
 - 2. Maintain Planned Resource Availability Database
 - 3. Maintain Preplanned Service Request Database and Scheduled Service Event Database
- D. Manage Performance
 - 3.c). (1) *long-term trend analysis*
 - 3.e) *send long-term planning input to Tech Ops*
- E. Manage Security
 - 1. Manage SN Security
 - 3. *Manage inter-system security*
- F. Manage Accounting
 - 1. Maintain SN Rate Database
 - 2. Maintain and report SN resource utilization
 - 3. *Report end-to-end Resource Utilization*

Italics denote inter-system control and co-ordination functions

Table 3.3-3

Both SNCA and SNCB perform all control functions listed in Table 3.3-1, albeit for a different set of SN resources and different set of external systems. There is limited coordination between SNCA and SNCB to resolve control of resources [such as DOD Lead Ranges (LRs) and GN] common to both subnets.

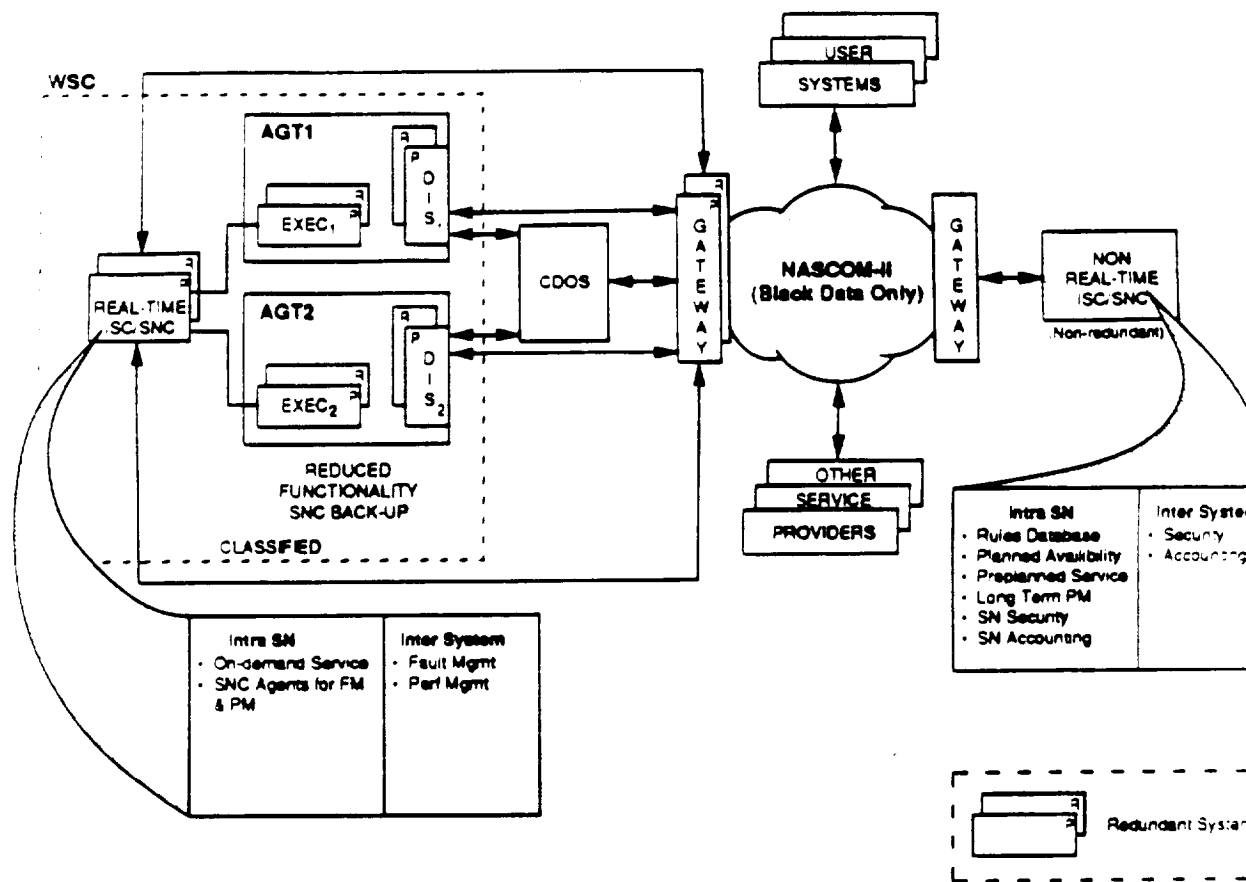


Figure 3.3-4: Architecture 2 - Real Time Function Partition

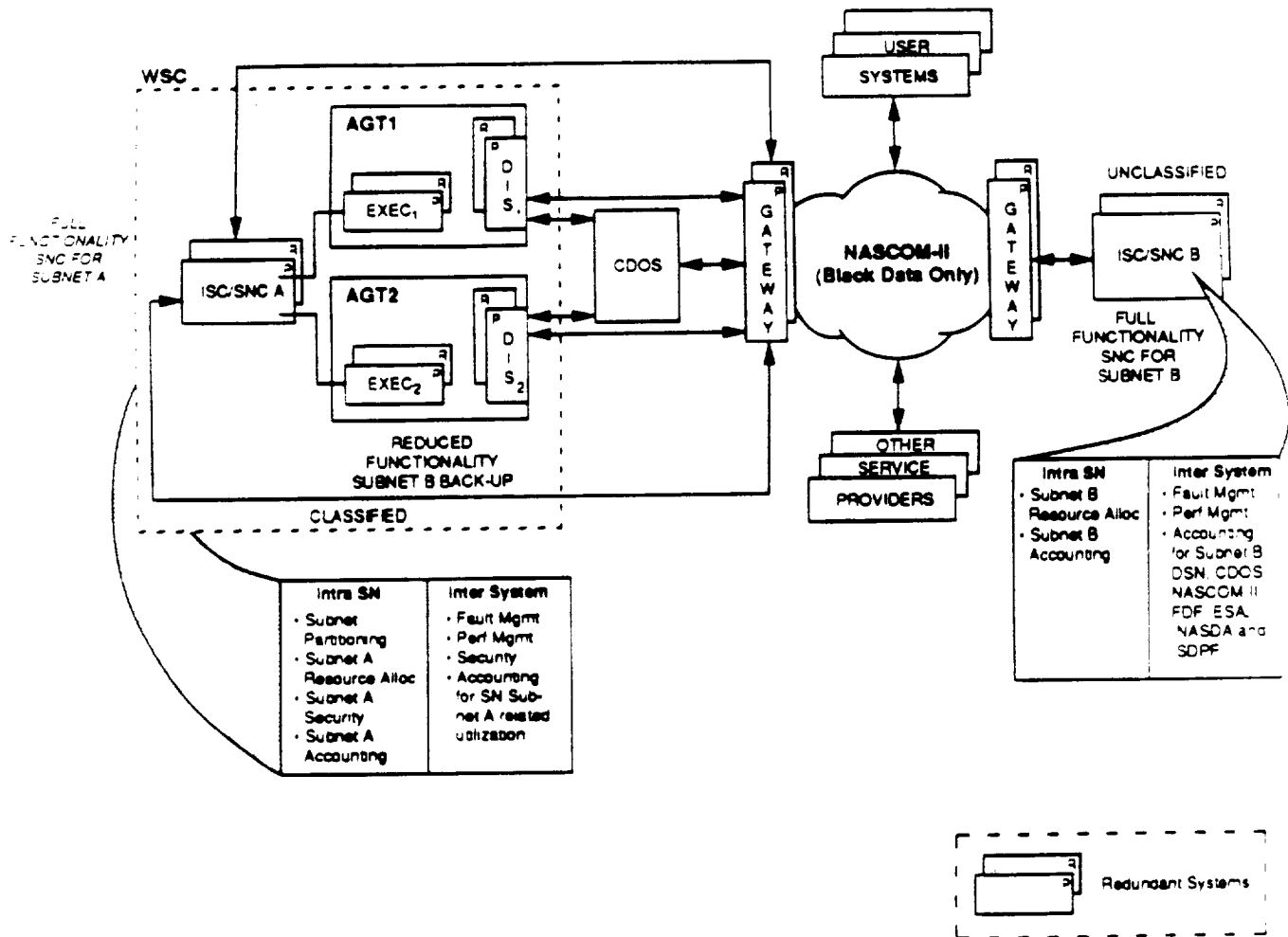


Figure 3.3-5: Architecture 3 - Dual Subnet Approach

This architecture provides limited location redundancy, in addition to control system redundancy for SNCA and SNCB. If the SNCB fails, subnet A can be reconfigured to include all or part of subnet B resources. SNCA can then operate as a full functionality backup SNC for the subnet B users. The reverse is not true, i.e., if SNCA fails, SNCB cannot take over subnet A control, since SNCB is unclassified. This is not a significant limitation, since complete failure of SNCA (located at WSC) is likely to occur in conjunction with complete failure of ATGT.

If the subnet partitioning is implemented, the ISC/SNCA can be integrated into ATGT. The pros and cons of such an integration are discussed in Section 3.3.7.

3.3.6 Standalone ISC Systems

The current NCC at the GSFC performs all intra-SN and inter-system control functions using some automated tools for intra-SN control and primarily manual procedures for inter-system controls. The key concept for ISC is monitor by exception. In this concept there is an operator assigned to every contact to ensure the quality of service provided by the SN is maintained at its high levels. However, the operator will be supporting multiple contacts concurrently. In order to do this, the operator must be supported with tools to:

- o generate alerts such that problems can be addressed in a timely fashion,
- o display the "big picture" of system status on demand.

This requires automation to collect the status data, process it, and generate user displays. For example, screens displaying a color coded fault diagnostic tree, as seen at the Naval Blossom Point Ground Station, would be very useful presentation formats enabling the operator to grasp the big picture quickly.

Many commercial networks use the concept of management of monitoring by exception. A typical system for a large corporation has one (or more) networks at each location. Each network is an autonomous entity with its own network manager. These networks are connected together by one (or more) long-haul backbone networks to form an enterprise system. The enterprise network control center(s) respond to exception reports, manage inter-system security and collect corporate-wide utilization information.

Functionally, the inter-system control/coordination (ISC) functions are independent of the intra-SN control functions and need not be performed by the same control system. The ISC functions can be performed by an autonomous system, which is separate from and independent of all other "systems" in the "System of Systems". Thus, the primary architectural alternatives are to continue with the current approach, Architecture 1, or utilize a standalone ISC, Architecture 4, as shown in Figure 3.3-6.

A standalone ISC system can be located anywhere and will normally communicate with the systems via NASCOM. If the level of redundancy (alternate routes) provided by NASCOM II does not meet the availability objectives, it could have its own back-up communication links for exchanging control information with the other systems, e.g., the public switched telephone network. As shown in Figure 3.3-6, the standalone ISC system can be implemented as a dual location redundant system (e.g. GSFC and WSC). If operating cost constraints do not allow a two location approach, the architecture can be modified to a one location redundant ISC system.

One of the strong attributes of this architecture is that it will not be dependent on or coupled to the design or development cycle of the systems to be controlled. Therefore, the standalone ISC system offers the maximum flexibility (ability to adapt the system to new user requirements and/or unanticipated modes of operations) and expandability (ability to accommodate increased workload).

The deployment cost of a standalone ISC system is estimated to be small compared to its development and operating costs. A standalone ISC can be implemented using SPARC 400 class workstations. The number of such workstations at an Inter-System Control Center (ISCC) will depend on the number of operators needed during peak traffic periods. This number is estimated to be considerably less than the number of operator consoles in the current NCC, since the proposed ISCC will be based on a monitor by exception concept. In this approach, operators will respond to exception reports provided by an automated process that is continually monitoring and analyzing status and providing progress reports.

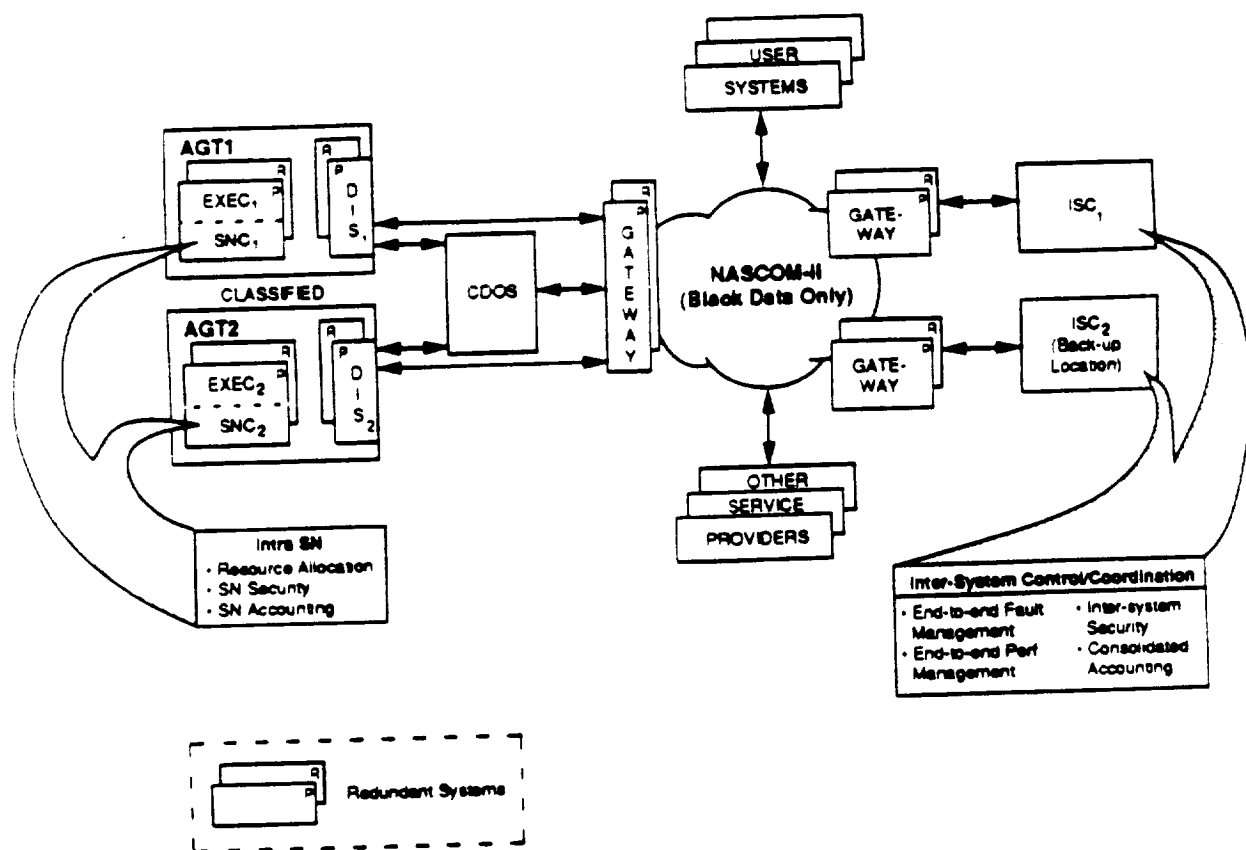


Figure 3.3-6: Architecture 4 - Stand Alone ISC

3.3.7 Integrated with SNC

Historically the NCC has been (and continues to be) responsible for resolving all inter-system interfaces. In addition, the SN is the primary "Space End System" and all space-ground information flows through it except during contingency services provided by Deep Space Network (DSN)/DOD LRs and international partners [European Space Agency (ESA) and Japanese Space Agency (NASDA)]. Therefore, a case can be made for the SN to be considered as more equal among equals. Other service providers (such as CDOS, NASCOM) exist to process or transport data in support of the SN services.

Integrating the SNC and ISC in an "Integrated SNC" may lower deployment and operational costs. At the same time, the coupling between SNC and ISC may bias the "Integrated SNC" in favor of the SN. This tends to limit the system flexibility and expandability.

3.3.7.1 Integrated Inter-system Control Architectures

As an alternative to the standalone ISC alternative, the ISC functions may be integrated into other systems. The ISC alternatives addressed in this section are:

- o Integrate with SNC (Architecture 1)
- o Integrate with ATGT
- o Integrate with CDOS.
- o Integrate with NASCOM II Network Management System (NMS).

These alternatives and their pros/cons are discussed below.

3.3.7.2 Integrated with ATGT

As discussed in Section 3.3.4, the real-time intraSNC functions may be integrated into the ATGT. An extension of this alternative is to enhance the ATGT to accommodate the ISC functions as well. The Automatic Data Processing (ADP) systems in the STGTs will be highly redundant and much more powerful compared to the ADP systems in the preSTGT systems. The STGT ADP systems (specifically EXEC ADPE) will be capable of performing control functions currently not performed by the WSGT or NGT, e.g. the pre-pass test of equipment to be used to support the contact. This could be extended to perform an end-to-end pre-pass test under the control of the ATGT.

This additional integration offers the benefit of simpler and more robust architecture for real-time intra-SN control functions. It can also lower the cost of the intra-SN control entity (SNC). It extends the benefits of the robust ATGT architecture to inter-system control functions as well. It has the disadvantage of coupling the ISC to the ATGT, which would require significant enhancements to the ATGT architecture and design. This tends to increase costs and limit the system flexibility/expandability.

3.3.7.3 Integrated with CDOS

The CDOS will provide an integrated mission data and operations management system for a large number of mission projects. The three primary services to be provided by CDOS are:

- o DDS,
- o Data Archive Service (DAS), and
- o CDOS Operations Management Service (COMS).

The CDOS DDS will process all CCSDS compliant data transported via the SN and the DSN. It will include a powerful Operations Management System. This system (COMS) could be enhanced to be made responsible for ISC functions as well.

Integrating the CDOS/COMS and ISC in an "Integrated COMS" may lower deployment and operational costs for CCSDS compliant data. However, integrating inter-system control for nonCCSDS data may require design changes to CDOS/COMS and result in higher design, deployment, and operational costs for the total CDOS. In addition, the close or tight coupling between COMS and ISC may bias the "Integrated NCC" in favor of the SN.

Also, the ISC will likely require the use of the composite SN schedule in its operation. Since the schedule may be classified, this may impose additional security requirements on CDOS.

3.3.7.4 Integrated with NASCOM II

The current NASCOM will be upgraded to a data driven NASCOM II. It will connect all NASA centers, data processing centers and international partners. Unlike the SN and the CDOS, NASCOM is not dedicated to transporting spacecraft data only. It will transport all other data to/from various NASA systems. It will include a powerful Long-haul Communications Network Management System (NASCOM II NMS).

The pros and cons of integrating the NMS and ISC in an "Integrated NMS" are similar to the pros and cons discussed earlier for Integrated SNC or COMS. The NASA NMS is primarily responsible for managing a ground-ground long-haul unclassified communications network. The ISC will be responsible for managing a variety of additional systems and interfaces. For example, security management, high-speed local/short-haul interfaces between SN/DIS and CDOS/SNIF. Therefore, an "Integrated NMS" may not be more cost effective. Furthermore, NASCOM II will be serving many other systems (that are not part of the "System of systems" addressed by this report). This will result in additional complexity, further increasing costs and risks and possibly compromise the level of attention provided to space-ground services.

3.4 Infrastructure Support

3.4.1 Introduction

In reviewing the POCC and NCC operations, the requirements for the end-to-end design of the SN related communications and information systems became apparent. The creation of a schedule is a complex co-operative process among groups (POCCs) desiring to use common SN resources with the SNC serving as arbitrator in the allocation of these resources. It is especially complex because the requirements are dynamic with changes being introduced throughout the planning cycle.

The process is currently much more complex than it needs to be with modern technology. As shown in Figure 3.4-1, the NCC data management is not integrated with the POCC systems in the sense that when data elements in one system change, they are not updated at the other. For example, in the Multisatellite Operations Control Center (MSOCC), a generic scheduling system is available for generating schedule requests, but the interface between it and the Mission Planning Terminal (MPT) is manual. When a reject message is returned by the SNC, the original request is not maintained for modification at the NCC, and the local database at the POCC is not updated.

To rectify these problems, it is envisioned that the SNC will be an integrated communications and information system with a major reduction in the manual co-ordination required to accommodate the inevitable conflicts and changes. The infrastructure necessary to support this environment involves **distributed data management, distributed work management, communications, automated interface management, information resource dictionary and computer security**. This infrastructure is described below.

3.4.2 Distributed Data Management

It is envisioned that the schedule will be maintained by a distributed data management system as depicted in Figure 3.4-2. The schedule database could also be implemented with a centralized approach. However, this would introduce considerable communications traffic between the POCC and SNC because the users require access to this information for their planning activities. For this discussion, the distributed approach is assumed.

In its planning the POCC scheduling system would generate a database of its schedule requests. From the instant that a schedule request is first submitted to the SNC until it is executed or deleted, it would be under the control of this system. The complete schedule will be maintained at the SNC, but the POCC specific elements of the schedule would also be stored locally. The POCC would be able to enter and delete requests, but only the SNC would be able to authorize them by allocating ATDRSS assets. The POCC would be able to modify requests providing they did not impact other users; otherwise the SNC would have to authorize the modification. With such a system, there would no confusion concerning the status of a schedule

CURRENT MSOCC

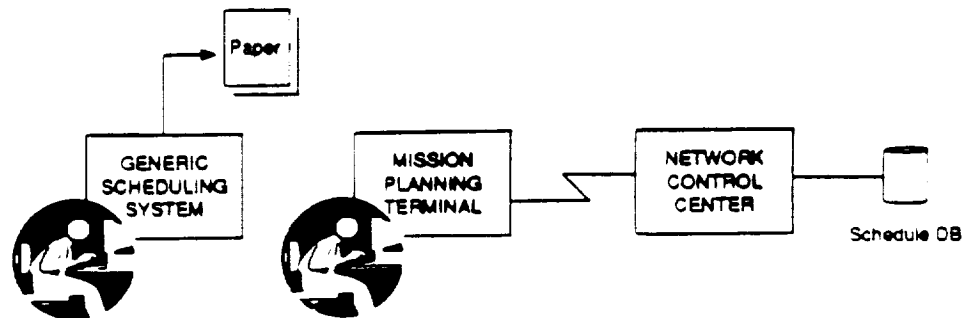


Figure 3.4-1: Current Approach

FUTURE

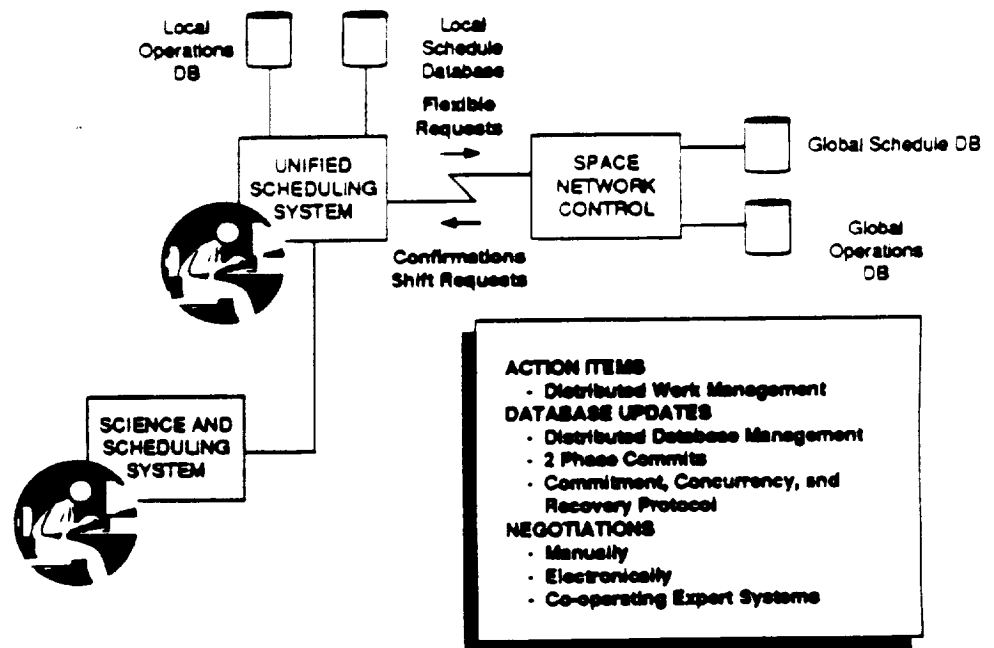


Figure 3.4-2: Recommended Approach

request. For example, when a user enters a modification to a schedule request in the local database, it would be:

- o modified locally and the update sent to the SNC providing the user had the authority to make the modification
- o set at a pending status and forwarded to the SNC for approval if the user didn't have approval; when the SNC (person or computer) evaluated the modification request, its status would be updated at the SNC and the local POCC.

To achieve the benefits of an integrated data management system, there are a number of complexities that must be handled. The major issue is how to accommodate changes in the schedule when multiple POCCs are affected.

Suppose a user POCC enters a new schedule request that can only be accommodated by moving the allocations of two other user POCCs. The SNC software would identify a shift request, and the SNC schedule database would be updated at the SNC reflecting that a change was in progress. Rather than have operators communicate using the telephone, the SNC would send shift requests via electronic mail to these POCCs tagged with a task to assess the impact and an action item to respond to the SNC. The management (creation, tracking, and updating) of the tasking and the action items is referred to as Distributed Work Management (DWM). By sending out action items and tasking, it is envisioned that the volume of verbal communications can be drastically reduced. DWM is viewed as an emerging technology as evidenced by the relevant products that have recently appeared in the Personal Computer (PC) marketplace such as Rhapsody (ATGT) or Notes (Lotus).

One of the key issues in generating the shift request is to what extent could the SNC could provide generic scheduling capability and take into account user spacecraft constraints. One approach is to have the SNC scheduling algorithm driven by time windows provided as input to the SNC by the user POCCs. In this case the translation from a generic scheduling input to a time window format is done in the POCC; the spacecraft constraints are implicitly specified via the windows. These windows would be defined in a sufficiently broad context to allow multiple orbits and inter-assignment constraints to be specified, e.g., a ten minute assignment between 2 and 2:30 or a ten minute assignment between 3 and 3:30. Alternatively, the SNC could be explicitly provided in a database of user spacecraft constraints as defined by the POCC to consider in generating the shift request. The major issues to be addressed are development and operational complexity of generic scheduling. From a development point, the issue is whether a sufficiently generic scheduling package could be developed that could accommodate all of the individual nuances and constraints of individual existing and planned spacecraft. If not, the SNC could provide a set of reusable software modules to the POCCs performing the basic generic scheduling functions. The POCCs would then tailor them to their specific needs. From an operational point of view, the issue is processing and managing the constraint database within the SNC. If the database is extremely large, it may be preferable to off-load the SNC and allocate the generic scheduling to the POCCs.

After the shift request is received at the POCC, the tasking to evaluate its impact be performed and a response generated to the action item. These actions may be performed by a person or a computer-based expert system. This may require the POCC to interact with the science user. Similar distributed data management, communications, and distributed work management technologies are applicable. If the POCCs agree to the shift request, then the POCC and SNC databases would be updated to reflect the shift. Otherwise, an alternative shift would have to be identified, the request put on hold, or the request rejected.

Modern distributed database systems equipped with two phase commit capability are able to support these distributed types of operations. Although current systems are largely vendor proprietary, standards are merging to support this capability. Therefore, this component should be obtained as an off-the-shelf item for integration into the SNC.

3.4.3 Communications

To provide the interaction between the SNC and the POCCs, an advanced communication system such as that defined by the ISO OSI protocol suite will be needed. Since

the government has heavily endorsed OSI by establishing Government OSI Profiles Interconnection (GOSIP), it is expected that OSI protocols will be used for SNC communications where practical. For the SNC, the major communications services envisioned for SNC are:

- o distributed transaction processing to support schedule distribution, and modification,
- o message handling to support electronic mail,
- o file transfer to transfer accounting and performance data,
- o graphics terminal emulation to access applications (this will allow the users to be remotely situated from the host computing systems), and
- o remote operations to support network management applications

With expected implementation of the SNC in the 1997 timeframe, the software implementing these services and the underlying communications protocols will be standardized and should be available off-the-shelf.

OSI protocols will provide new options for implementation of encryption to provide security functions. Standard protocols for implementing these functions are currently being developed by National Institute of Standards and Technology (NIST) as part of the Secure Data Network System (SDNS). Rather than use link encryption as currently done in NASCOM, encryption may be performed at higher layers. Since NASCOM II will be a mesh network, all communications between the WSC and GSFC may not be provided by a point-to-point link. Thus, higher level encryption may be preferable to link encryption. As shown in Figure 3.4-3, one alternative is to perform encryption between the network and transport layers; the two protocols for supporting this alternative are:

- o SP3 - encryption on a source-destination basis,
- o SP4 - encryption based on a transport connection basis.

As illustrated in the figure, these protocols reside above the network layer to be supported by NASCOM II. Therefore, the SNC may have to provide the SDNS Key Management Center. Although this imposes an additional security activity, it should be implemented using off-the-shelf components in the 1997 time-frame.

Another SDNS alternative being developed by NIST is encryption at layer 6 or 7 for electronic mail applications. Use of this option may also require the SNC to support a Key Management Center.

3.4.4 Automated Interface Management

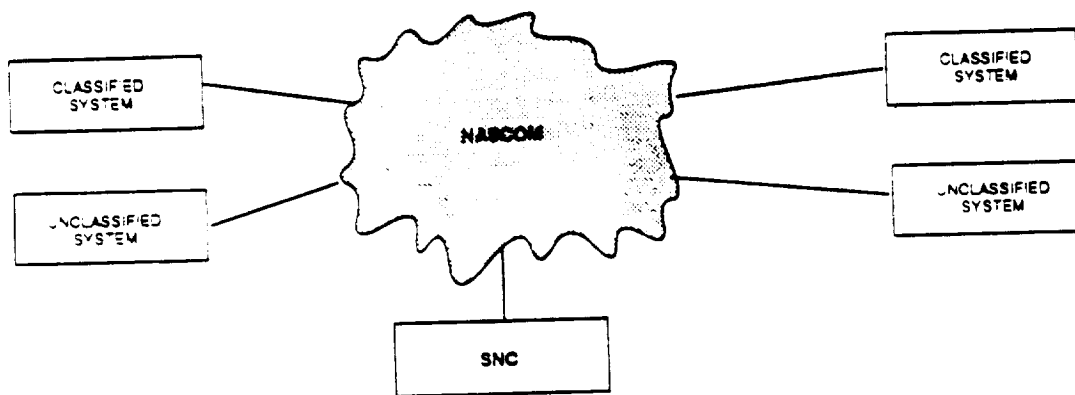
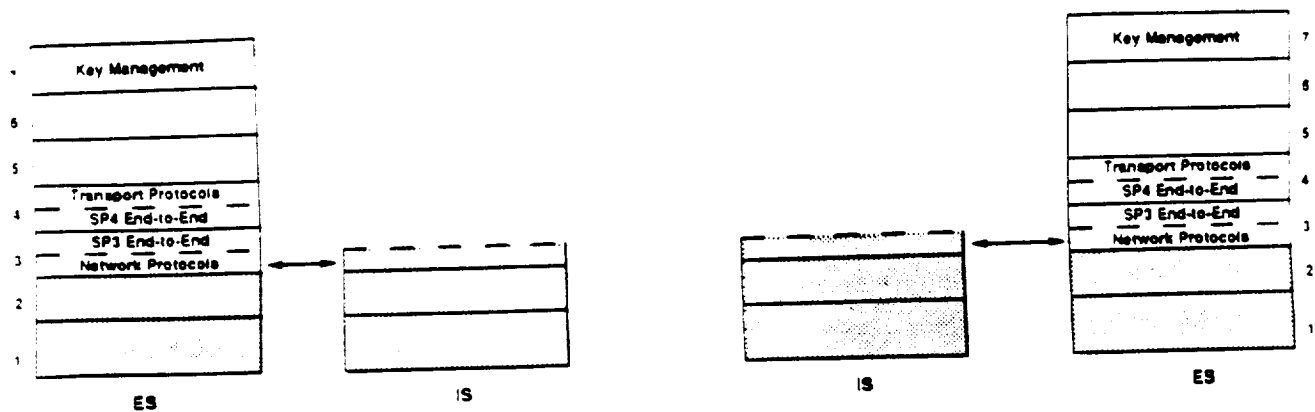
Considering the broad variety of interfaces needed to support the communications services described above, management of these interfaces will be complex. However, it will be greatly simplified by the use of utilities that have evolved to support OSI applications. The primary utility of interest is the ASN.1 compiler discussed below.

In OSI the applications layer interface is described using a special language referred to as the Abstract Syntax Notation One (ASN.1) (ISO 8824). Using a basic set of constructs (integer, character) and a composite set of constructs (sequences), ASN.1 provides a very general message definition capability. Messages will then be encoded using a Type-Length-Value (T-L-V) algorithms referred to as the Basic Encoding Rules (ISO 8825).

ASN.1 compilers provide the capability to automatically generate the encode and decode software from the ASN.1 specification of the application interface. The encode software accepts as input a data structure corresponding to the message and generates the T-L-V encoding. Conversely, the decoding software accepts a T-L-V byte stream as input and instantiates the data structure corresponding to the message. Thus with the use of ASN.1, the communications software is easily changed when the interface is modified.

3.4.5 Information Resource Dictionary

In the case we suggest that a set of institutional standards for defining and exchanging planning and scheduling data be developed and adopted. Such standards would enable planning and scheduling applications to communicate at the semantic level (subject to security and privacy constraints).



LEGEND:

ES - End System
IS - Intermediate System

Figure 3.4-3: SECURE NETWORK TOPOLOGY

Although the existing NCC already has a defined standard interface with the POCCs, it is recommended that the SNC planning and scheduling system be based on a more general standard capable of unifying all of the interacting systems. Such a standard would lay the foundation for a "seamless" planning and scheduling network, within the system-of-systems, capable of rapidly responding to non-nominal situations with maximally productive schedules.

An appropriate standard already exists for this purpose, called the Information Resource Dictionary Standard (IRDS). The IRDS is a four-layer model for defining data. At the top level, data are modeled using three basic constructs: entity types, relationship types, and attribute types. The Layer 2 model, or *data definition schema*, is created by defining specific entity types, relationship types, and attribute types describing the general domain of planning and scheduling.

A key aspect of the technical approach is the use of a standard notation for expressing potentially complex time and quantity relationships and constraints. This notation should provide, as a minimum, all of the expressive capability provided by the Flexible Envelope Request Notation (FERN) developed by NASA.

Given the standard layer 2 model, each mission and facility would thus construct its own layer 3 model in accordance with the standard. In the process, all of the static constraints that exist between instances of these activities would be captured for general use. The fourth layer of the IRDS model would then represent the specific plans, schedules, and ephemeral data produced or used by the missions and facilities in the course of flight operations.

The major benefit of adhering to the IRDS standard is that commercial off-the-shelf (COTS) software could be used to create and manage the data dictionaries, as well as providing a standard interface for the reuse of software components across multiple systems.

3.4.6 Computer Security

As discussed in Section 3.2.2, one alternative for operation of SN scheduling in the unclassified mode is to allow the users to work together to generate the schedule. This imposes an additional requirement for applications interoperability among the POCCs such that their mission planning and scheduling data can be interchanged.

The security infrastructure definition for the SNC is based on the ability to either partition the classified data from unclassified data into separate processing environments, or to segregate the majority of the security functionality into a highly trusted subsystem. The three basic computer security alternatives for the SNC are a Multi-level secure (MLS) architecture, a data partitioned architecture, and a functionally partitioned architecture. Each of these alternatives is presented in the following subsections.

3.4.6.1 Multi-level Secure Architecture

The MLS architecture as depicted in Figure 3.4.6-1 presupposes that the evolution of "trusted computing systems" has evolved to the point that a B2-class system (as defined in DOD 5200.28-STD) can be established as the processing core of the SNC. The security management functions provide the foundation in which the remainder of the (less trusted) SNC functions operate.

The majority of the security functionality is allocated to a secure COTS operating system or security kernel, including, access control, auditing, and interfacing to the communications media. A secure data base management system may be used to control and audit access attempts to the data it manages (e.g., the schedule). A set of manual procedures would be used to maintain the integrity of the security functions such as establishing login ids, maintaining security designation(s) of communication channels, reviewing audit trail data.

The Multi-level Secure architecture would be used as the infrastructure for either the Integrated SNC or the Real-time Partitioned architectural alternatives presented as Alternatives 1 and 2 respectively, in Section 3.3.

3.4.6.2 Functionally Partitioned Architecture

The Functionally Partitioned Architecture as depicted in Figure 3.4.6-2 is based on the classical "system-high" mode of computer system operation. In this case, the majority of the

security functions are isolated into a separate processor referred to as the "guard processor". The guard processor is responsible for authenticating and auditing user access as well as downgrading data that can be sent to unclassified systems. The security data base for the most part, is maintained by the guard processor. All interfaces to unclassified systems are required to access the SNC only through the guard. Other classified systems may be allowed to bypass the guard processor, if no violation of the security policy is possible.

The Functionally Partitioned architecture would be used as the infrastructure for either the Integrated SNC or the Real-time Partitioned architectural alternatives presented as Alternatives 1 and 2 respectively in Section 3.3.

3.4.6.3 Data Partitioned Architecture

The Data Partitioned Architecture as depicted in Figure 3.4.6-3 attempts to isolate the classified data and processing performed by SNC from the unclassified portion. This partitioning is based on the premise that it is possible to partition the SN resources so that the user POCC's can gain less restricted access to a portion of the SNC.

A portion of the security functionality is allocated to both the classified and unclassified subsystems. These subsystems are still responsible for authenticating user access to their respective data stores. Because the classified subsystem is not directly communicating with the unclassified user POCC's, the overall level of security functionality is somewhat reduced over the MLS system above.

To avoid possible security policy violations, however, the classified and unclassified portions of the SNC would still need to interact in a very restricted fashion, probably through a small security guard processor. This interaction is needed to pass distributed information between the two processors such as accounting, performance and security audit data. The majority of the security functionality is therefore allocated to the guard processor to ensure that classified data is not released to the unclassified subsystem. Manual procedures are still used to establish and maintain the security data base needed by the system.

The Data Partitioned architecture would primarily be used as the infrastructure for the Classified Partitioned architectural alternative presented as Alternative 3 in Section 3.3.

3.4.6.4 Conclusions

The analysis results of the security infrastructure indicate that the Multi-level Security (MLS) alternative architecture has several major drawbacks including very high development and maintenance costs, lower performance and reduced flexibility over the other options. For that reason, this alternative has been dismissed as a feasible approach. Therefore, if the selected SNC architecture is either Alternative 1 (Integrated SNC) or Alternative 3 (Real-time Partitioned), the Functionally Partitioned security architecture is recommended. The Data Partitioned security architecture is recommended if Alternative 2 (Classified Partitioned) is the selected SNC architecture.

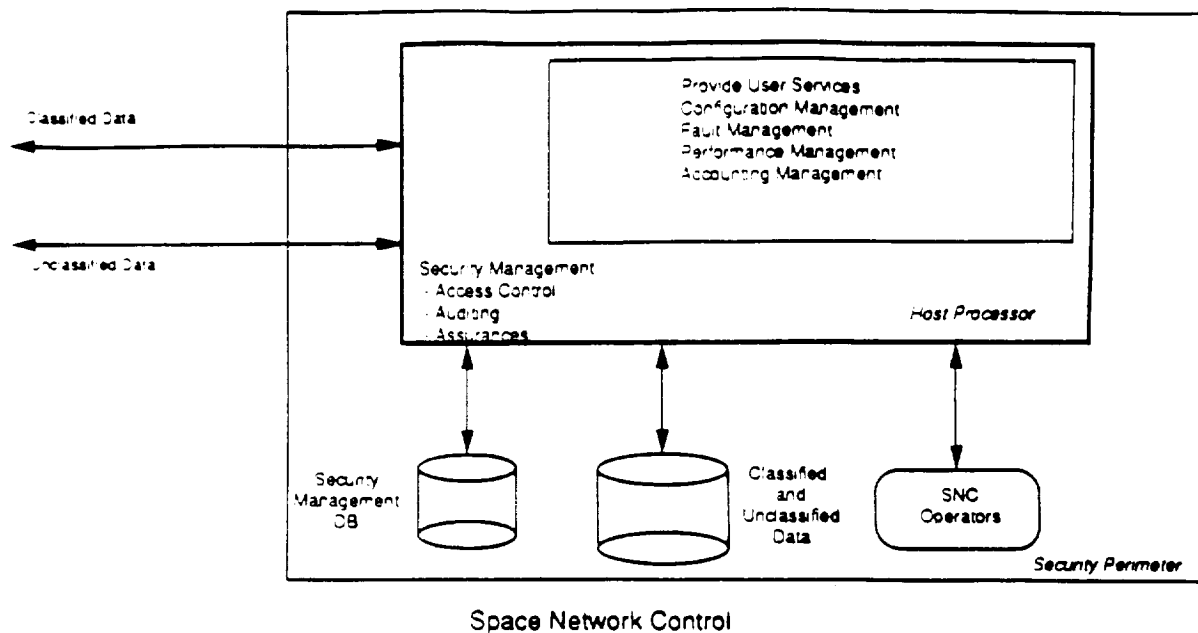


Figure 3.4.6-1: MULTI-LEVEL SYSTEM ARCHITECTURE

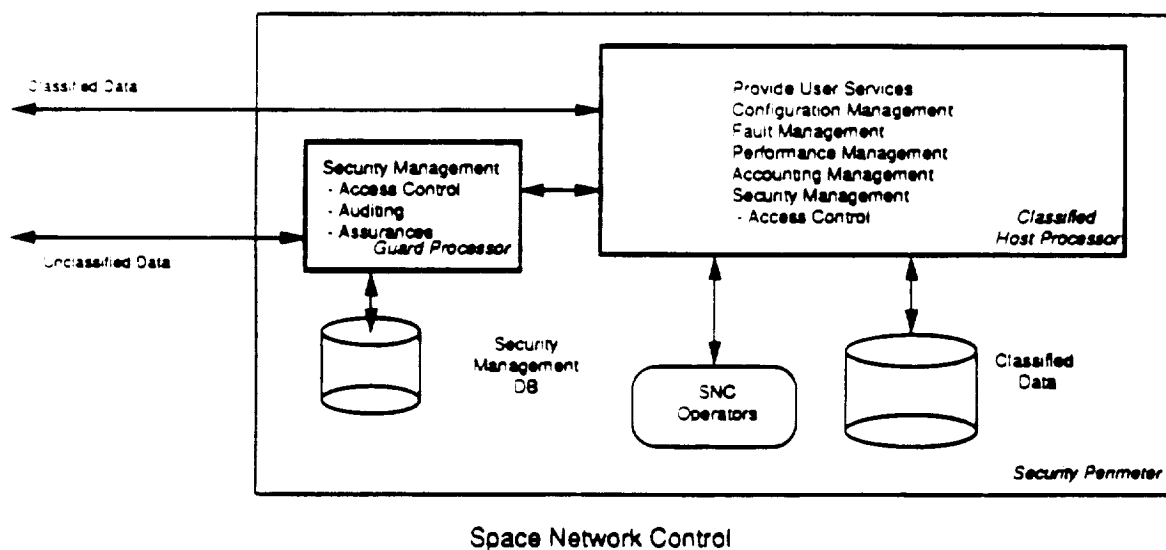


Figure 3.4.6-2: FUNCTIONALLY PARTITIONED ARCHITECTURE

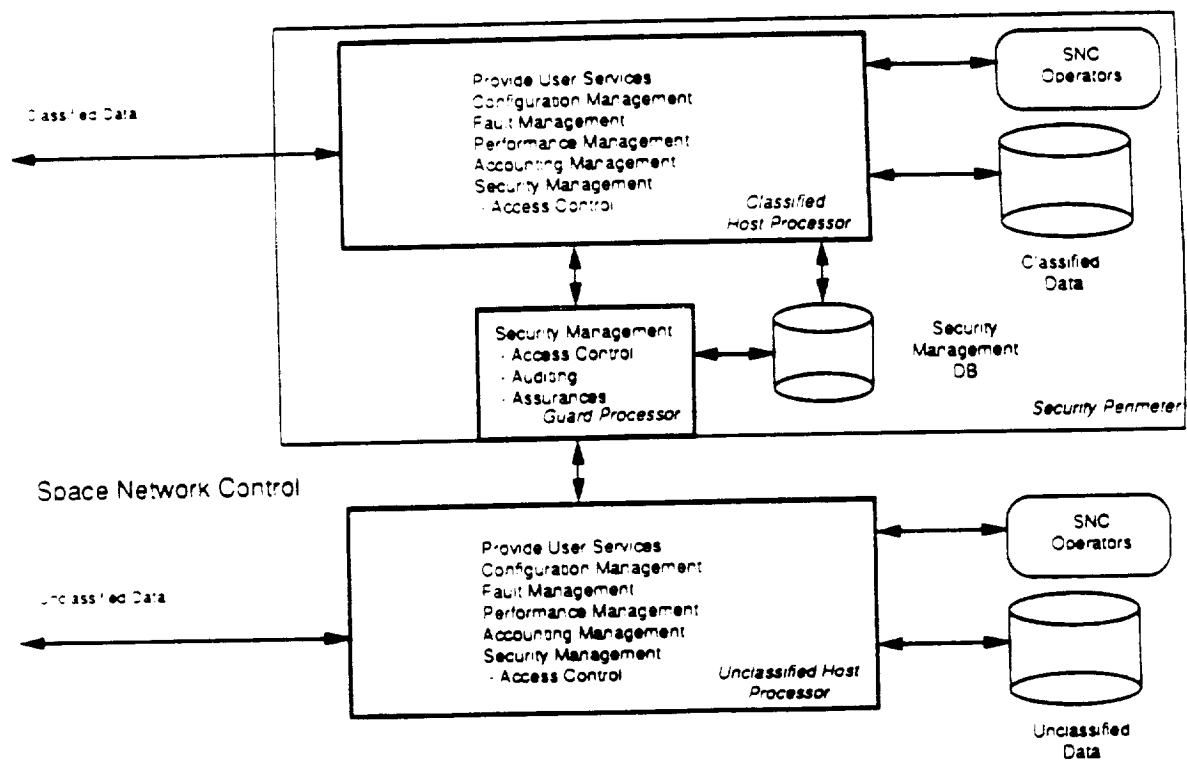


Figure 3.4.6-3: DATA PARTITIONED ARCHITECTURE

4. Tradeoffs

In this section the alternatives formulated in Section 3 are compared against a set of evaluation criteria in order to determine the tradeoffs. The evaluation criteria are defined in Section 4.1.

In Sections 4.2 to 4.6 the major alternatives formulated in Section 3 are evaluated. As shown in Table 4-1, these alternatives consist of:

- o Resource Allocation
- o Resource Partitions
- o Real-time and Non-Real-Time Partitions
- o Integration of the Space Network Control(SNC) with Advanced Tracking and Delay Relay Satellite System (ATGT)
- o Introduction of Automated Inter-System Control (ISC)

This comparison is largely qualitative although performance of the resource allocation algorithms is evaluated quantitatively. The quantitative results were generated using the National aeronautics and Space Administration (NASA) Network Planning and Analysis System (NPAS) model as well as a CTA INCORPORATED (CTA) developed model for real-time demand access.

4.1 Definition of Evaluation Criteria

To discriminate between alternatives, a set of evaluation criteria was defined and applied to the alternatives. They are as follows:

- Capability - The degree to which user needs and service requirements are met by the conceptual alternative under normal fault conditions,
- Performance - The ability of the conceptual alternative to meet performance goals of probability of obtaining service, network utilization, and user responsiveness,
- Risk - The schedule, technical, and cost uncertainty associated with the conceptual alternative,
- Flexibility - The ability of the conceptual alternative to adapt to new user requirements or unanticipated modes of operation,
- Expandability - The ability of the conceptual alternative to accommodate an increased workload of schedule requests and users,
- Cost - The life cycle costs required to implement the conceptual alternative.
- Ease of Use - The ease of use experienced by users of the conceptual alternative; the amount of training required to bring users to a level of proficiency.

4.2 Capability

Capability is the degree to which user needs and service requirements are met by the conceptual alternative under normal operating conditions. As described in Section 2.3, there are several modes of mission operations with differing needs for Space Network (SN) services. These modes are:

- o Mode 1 - Single Event Operations
- o Mode 2 - Emergency Operations
- o Mode 3 - Target of Opportunity Operations
- o Mode 4 - Aperiodic Operations
- o Mode 5 - Periodic Operations

How a conceptual alternative will respond to each of these modes is of prime importance in evaluating the alternative, especially in the case of the resource allocation alternatives. The evaluation that follows, therefore, is largely concerned with how the alternatives respond to the needs of a mission operating in each of these modes.

4.2.1 Resource Allocation

4.2.1.1 Demand Access

The demand scheduling approach provides the best response to emergency and target of opportunity requirements. If demand scheduling were available with a high probability of successfully scheduling a contact, perhaps more missions would be developed to operate in a target of opportunity mode. A benefit of this could be a decrease in total data traffic. This decrease could come about if missions could achieve substantially the same results from collecting data at carefully selected times instead of operating in periodic or aperiodic mode, collecting data continuously.

This scheduling approach would be robust in the event of failure of SN resources, either temporary or permanent. Overall performance would decrease, in proportion to the significance of the lost resources, but no last minute rescheduling would be needed, and no long-planned operations would be cancelled.

Demand scheduling may not be acceptable to missions with known needs for frequent aperiodic or periodic sessions, unless blocking probability can be kept very low. Demand access is not suitable for single event support for which a guaranteed contact at a specific time is required, unless some sort of priority with preemption scheme is used to ensure access.

The workload may have peaks and valleys, leading to varying demand on SNC staff. Uncertainty as to whether a particular attempt to schedule a contact will succeed could make personnel scheduling in the Payload Operations Control Center (POCC) difficult as well.

4.2.1.2 Pre-Planned Allocation

Pre-planned allocation is desirable for missions with well-defined needs that seldom change. It permits advanced planning of staffing for both mission operations and for SN. Operations tasks that require advance preparation are best served by an approach that permits advanced reservation of resources.

Emergency and target of opportunity operations, on the other hand, are not handled well by a pre-planned system, because a "frozen" schedule makes it hard to accommodate pop-up requests to handle these requirements. This is less of a disadvantage if the "liquid" and "slushy" states predominate. If the schedule is frozen too far in advance, even missions that operate in Mode 4 (Aperiodic) may find it difficult to maximize data return because a schedule that seemed optimal at the time it was formulated may not match the actual downlink requirements. This effect may get worse with the increasing use of intelligent instruments that adapt their data collection rates and on-board processing to suit observations.

Failure of SN resources would have serious repercussions on a pre-planned schedule in the short term. Because there might not be time to replan, a failure could result in a high priority session being cancelled, while a lower priority mission is unaffected. In the long run, the pre-planned approach would take permanent outages into account in developing schedules.

4.2.1.3 Short Term Scheduling

The short term scheduling approach combines features, both good and bad, of the demand access and pre-planned approaches. It represents a compromise approach to allocation of SN resources.

For example, servicing of periodic and aperiodic needs that are known well in advance is fulfilled better by short term scheduling than by demand access, but not as well as by a pre-planned approach. Short term scheduling may not allow adequate preparation time prior to a contact, particularly if change or shift requests are needed. The difficulty of personnel scheduling when the ability to schedule a contact is uncertain is nearly as bad as for demand access. Also, scheduling of service for single events (mode 1 operations) is fulfilled better by short term scheduling than demand access, but not as well as pre-planned. Unless service is virtually guaranteed by a high priority, the greater advanced scheduling of the pre-planned approach would be preferred for launch, orbit operations, and other such operational activities.

Servicing of emergency operations and target of opportunity requests are better served by short term scheduling than by the pre-planned approach, but not as well as by demand access. In both of the above situations, the ability to schedule a contact shortly before the desired start time makes short-term scheduling preferable to the pre-planned approach. But there is still a freeze point hours before the event, thus demand access would respond better to these needs.

In the event of failure of SN resources, the short term approach would react much like the pre-planned approach.

4.2.1.4 Hybrid Allocation

Hybrid allocation provides the best of both the demand and pre-planned allocation approaches, accommodating all modes of mission operation, but it avoids their greatest disadvantages. The hybrid approach, by providing demand access to some resources, can satisfy requirements for emergency and target of opportunity operations. But since this approach also provides for advanced scheduling of contacts, single event, aperiodic, and periodic operations can be handled as well. A great deal of excess capacity may be needed to assure support for last-minute requests after allocating resources in advance for missions that need to schedule services in advance.

Failure of SN resources would result in disturbance of the pre-allocated part of the schedule, as in the pre-planned approach, and an increased probability of blocking for demand requests, as in the demand access approach. An advantage of the hybrid approach is that a high priority mission that had a pre-planned contact bumped because of a resource failure would be able to make a demand request, and thus might be able to conduct the planned contact, despite the failure.

4.2.2 Resource Partitions

A potential advantage of this partitioning is that the resource allocation approach could be optimized to each user group or partition. This could mean a different approach for each partition, or simply a variation in the parameters of a single approach. For example, in the demand access approach, the priority scheme or method used to handle blocked requests might vary from partition to partition. Another advantage is that the schedule for the dedicated resources of each partition could be published, making it easier to make requests, and decreasing the amount of conflict resolution processing by both users and SNC.

Separation of manned from unmanned missions would relieve unmanned users from the uncertainty of scheduling at planned shuttle launch times and during shuttle missions. Gaining this relief comes at the expense of removing enough resources from the pool to accommodate the high priority shuttle requirements, whether they turn out to be needed or not.

A possible disadvantage is that there could be an imbalance in allocation of resources to the groups, leading to frequent blocking for users in one group, while another group's resources are under-utilized. Despite the "spillover" concept, scheduling of another group's resources could be much less satisfactory than using the resources of one's own group (e.g., due to lower priority, or restriction to last-minute demand access). Such an imbalance could be caused (or made worse) by failure of resources allocated to one partition.

4.2.3 Real-time vs. Non-real-time Functional Partition

Not Applicable.

4.2.4 Integration of SNC with ATGT

These options have little impact on capability as seen by users, except that some options may make it easier to operate with resource partitions.

4.2.5 ISC

The end-to-end capabilities of a Automated ISC could be of benefit to users because more effective use of resources could result. This would include not only SN resources but also those of the user missions and institutional facilities [such as Customer Data Operations Services

(CDOS) and NASA Communications Network (NASCOM)]. An integrated approach to management and control of ALL resources needed to perform a task would simplify operations and could lead to lower operating costs.

The system of systems challenge, however, is formidable. Retrofitting the Automated ISC approach into missions that are being developed now [e.g., Earth Observing System (EOS)] assuming NCC-like operations may be difficult -- technically, financially, and organizationally. It would be necessary to create an evolutionary approach that requires new missions or institutional upgrades, while allowing older missions to run in Network Control Center (NCC) mode.

4.3 Relative Cost Analysis and Rankings

Life cycle cost consists of three primary elements:

- o Development Costs (DEV) - this includes the cost of developing software, procedures, interface definitions and development system.
- o Deployment Costs (DEPLOY) - this includes the cost of acquisition, installation, training for all operational hardware, software, communication services for SNC and User Interface systems (if applicable).
- o Operational Costs (OPER) - this includes the facility, personnel, and computer system and maintenance/service costs.

All cost analysis is relative ranking, with 1 being the lowest cost alternative and 5 being the highest. At an architectural level, absolute or relative costs cannot be estimated accurately, since cost is a function of many design and implementation decisions beyond the scope of this study.

4.3.1 Resource Allocation

The relative cost rankings for the resource allocation alternatives defined earlier in section 3.2 are as follows:

	DEV	DEPLOY	OPER
Current NCC approach	2	2	4
Pre-planned fixed	3	3	2
Pre-planned fluid	4	4	3
Short-term (2-6 hrs)	3	5	1
On-demand (5-15 min)	1	1	1
Hybrid	5	4	2
(fluid+stretched on-demand)			

Development cost ranking

The simplest allocation algorithm is the on-demand alternative. With this approach, service events are scheduled as soon as the service requests are received. There is no need to keep track of and schedule events at a future time. In addition, with the on-demand approach there is no need for a classified schedule or distributed data/work management. Therefore the on-demand approach has been assigned the rank order 1.

If the current NCC approach is continued in the Advanced TDRSS Ground Terminal (ATGT) era, it will require complete redesign and development effort to support increased SN capacity/channel types. The current NCC approach supports a combination of pre-planned fixed and on-demand resource allocation algorithms with an emphasis on the pre-planned approach. All other alternatives (except on-demand) will require some development effort to support the "User System Information Interface" using distributed data and work management as described in Section 3.4. Therefore, the current NCC approach has been assigned the rank order 2.

Among the remaining alternatives, the development costs will be a direct function of automated allocation algorithms. The pre-planned fixed and short term alternatives are more complex than the on-demand approach, since both support scheduling of future events. These alternatives are less complex than the pre-planned fluid alternative, which breaks down the scheduling process in three phases (liquid, slushy and frozen) and may use several allocation

strategies (demand leveling, window pruning, and constraint relaxation). The complexity of the pre-planned and short term alternatives were judged to be comparable. Therefore both have been assigned a ranking of 3 and the preplanned fluid has been assigned a ranking of 4.

The hybrid approach is a combination of pre-planned fluid and modified (stretched) on-demand approach. Therefore, it was judged to require maximum development effort and has been assigned a ranking of 5.

Deployment cost ranking

The deployment cost of hardware/software will be a function of the algorithm complexity and frequency of rescheduling, while the training costs will be a function of the user interface complexity. The on-demand alternative is the simplest and therefore has been assigned a ranking of 1. If the current NCC approach is continued, there will be no need to deploy distributed data or work management systems. Therefore, it has been assigned the rank order 2.

The short-term alternative requires rapid rescheduling as each service request is received. It cannot spread the processing over a period of time, as is the case with pre-planned alternatives. Therefore, the short-term alternative was judged to require the maximum processing power. The training cost for this alternative is more than the on-demand approach but less than all other automated approaches. The higher cost of processing system was judged to exceed any savings in the training costs, and therefore the short-term alternative has been assigned the highest cost ranking of 5.

The hardware required for the pre-planned fixed alternative would be more than that required for the current NCC approach, to support distributed data and work management. The pre-planned fixed algorithms are simpler compared to the pre-planned fluid and the frequency of scheduling runs is also less. Therefore, the pre-planned fixed has been ranked as 3, while the pre-planned fluid has been ranked as 4.

The hybrid approach is a combination of pre-planned fluid and modified (stretched) on-demand approach. Therefore the deployment costs for this alternative is comparable to the deployment cost for the higher of the two alternatives, i.e., pre-planned fluid.

Operational cost ranking

Operational costs are primarily a function of the amount of negotiations, complexity of user interface, communications costs and hardware/software maintenance costs.

The current NCC approach was judged to require maximum negotiations and the most complex user-interface. Therefore, it has been assigned the highest cost ranking of 4. The user interface for the on-demand and the short-term alternatives are simplest and of comparable complexity. These also require the least amount of negotiations and therefore have been assigned a ranking of 1.

The pre-planned fluid would require more negotiation compared to the on-demand or short-term approach, since it schedules future events. For the same reason, its user interface would be more complex. However, it would require less negotiation compared to the pre-planned fixed approach. The user interface complexity for both pre-planned approaches is comparable. Therefore, the pre-planned fluid and fixed have been ranked as 2 and 3 respectively.

4.3.2 Resource Partitions

The relative cost rankings for the resource partitioning alternatives described earlier in section 3.3.1.2 are as follows:

	DEV	DEPLOY	OPER
Current NCC approach	1	1	2
[Unpartitioned Classified SNC at Goddard Space Flight Center (GSFC)]			
Classified and Unclassified at GSFC	2	2	2
Classified at White Sands Complex (WSC), Unclassified at GSFC	2	2 or 3	1

Development cost ranking

Supporting partitioning of resources would incur additional development costs compared to a system without any resource partitioning. However, the specifics of the partitioning approach is not expected to impact the development costs, i.e., development costs for the two partitioning alternatives was judged to be comparable.

Deployment cost ranking

The deployment costs for the partitioned alternatives were judged to be higher than the cost of unpartitioned SNC system due to the cost of deploying two separate systems. We have assumed that both GSFC and WSC already have adequate classified facilities. If the ATGT classified facilities at WSC are not adequate to house the classified SNC system, the deployment cost ranking for the last alternative will change to 3.

Operational cost ranking

The unpartitioned alternative and the partitioned alternative with a classified system at GSFC will both require continued operation of a classified facility at GSFC and encrypted communication links between WSC and GSFC. Operating classified facilities and links is quite expensive. In general the total life-cycle cost (development, deployment and operation over life cycle) of a classified system is several times higher compared to the life cycle cost of an equivalent unclassified system. In other words, both the unpartitioned and the partitioned alternative with a classified system at GSFC will incur higher operational costs. The operational cost of these two alternatives were judged to be comparable.

4.3.3 Real-time and non real-time partitions

The cost tradeoffs of partitioning the SNC system based on real-time function partitions are as follows:

	DEV	DEPLOY	OPER
No partitions (all functions at GSFC)	1	2	2
Real-time functions at WSC and non real-time at GSFC	1	1	1

Development cost ranking

In general development costs for non real-time application are higher on a real-time system. Therefore, even if the real-time functions were kept at GSFC, the GSFC SNC system will consist of real-time and non real-time systems connected via a Local Area Network (LAN). In other words, the primary difference between the two alternatives is whether the real-time and non real-time systems are collocated (i.e. connected by a LAN) or not (i.e. connected by NASCOM II). This difference will not impact the development costs.

Deployment cost ranking

Keeping all SNC functions at GSFC will require additional communications equipment to support real-time control dataflow between WSC and GSFC. Therefore this alternative has been ranked as 2.

Operational cost ranking

Keeping all SNC functions at GSFC will incur additional communications costs to support real-time control dataflow between WSC and GSFC. In addition, there may be some savings in personnel by using the same people to monitor real-time events for the ATGT and SNC systems at WSC. Therefore the first alternative (keep all SNC functions at GSFC) has been ranked as 2.

4.3.4 Integration of Real-time SNC with ATGT

In the previous section, we showed that moving the real-time SNC functions to WSC offers cost savings. In this section, we analyze the cost impact of integrating the real-time SNC functions with the real-time ATGT system. The alternatives analyzed are:

- o real-time SNC system at GSFC,
- o real-time SNC system at WSC, and
- o real-time SNC functions integrated with ATGT.

	DEV	DEPLOY	OPER
No Integration			
real-time SNC at GSFC	1	3	3
real-time SNC at WSC	1	2	2
Real-time SNC integrated with ATGT	2	1	1

Development cost ranking

The location (GSFC or WSC) of the real-time SNC system does not change the development costs, i.e., the development costs of the first two alternatives are comparable. The two ATGT systems (AGT1 and AGT2) operate as independent systems. These systems have not been designed individually to be able to control all ATDRSS and ATGT resources concurrently. The real-time SNC functions will control all ATDRSS and ATGT resources as one SN subsystem. The cost of modifying the AGT1 and AGT2 subsystems to operate as one subsystem were judged to be considerable and therefore the integrated ATGT alternative has been ranked as 2.

Deployment cost ranking

It has been assumed that the ATGT computer/communication complex will not have excess capacity to support real-time SNC functions. Furthermore, it has been assumed that this capacity can be increased by adding or upgrading hardware without introducing new types of computer systems. The incremental cost for capacity increase was judged to be lower than the cost of deploying a separate (and most likely different type) real-time SNC system. Therefore, the integrated ATGT alternative has been ranked as 1.

A separate real-time SNC at GSFC will cost more than a separate real-time SNC at WSC due to the cost of additional communications equipment to support real-time control dataflow between WSC and GSFC. Therefore the real-time SNC at WSC and real-time SNC at GSFC have been ranked as 2 and 3 respectively.

Operational cost ranking

Keeping all SNC functions at GSFC will incur additional communications costs to support real-time control dataflow between WSC and GSFC. Therefore operational costs for real-time SNC at GSFC will be higher compared to the other two alternatives. Among the other two alternatives, the operational costs for a separate real-time SNC at WSC will be higher due to the maintenance and support costs of a separate (and most likely different type of) system. Therefore the alternatives have been ranked as 1 for integrated real-time SNC, 2 for separate real-time SNC at WSC and 3 for real-time SNC at GSFC.

4.3.5 ISC System

The architectural alternatives for the ISC function are:

- o continue the current NCC manual approach using multiple 2-way voice conversations,
- o a standalone automated ISC system at GSFC or WSC,
- o integrate ISC with SNC at GSFC or WSC,
- o integrate ISC with CDOS/CDOS Operations Management Services (COMS).

As indicated earlier in section 3.3, integrating NASCOM/Network Management System (NMS) and the ISC system is not a practical alternative. The relative cost rankings for the four alternatives are as shown below:

	DEV	DEPLOY	OPER
Current NCC approach	1	1	4
Stand-alone @ GSFC or WSC	2	3	3
Integrated with SNC	2	2	1
Integrated w CDOS/COMS	3	4	2

Development cost ranking

Continuing the current NCC approach would not require development of a highly automated ISC system, i.e., it is the lowest development cost alternative.

Among the three automated alternatives, integration with the CDOS/COMS would be the most expensive alternative due to the extra costs incurred in adding the support for real-time and nonCCSDS payload data to CDOS/COMS. The development costs for the ISC functions (only) for the standalone automated ISC system and an integrated SNC/ISC system will be comparable. Therefore these two alternatives have been ranked as 2 and the integrated CDOS/COMS alternative has been ranked as 3.

Deployment cost ranking

Continuing the current NCC approach would not require deployment of an automated ISC system so it is the lowest deployment cost alternative. The deployment cost for a standalone ISC system will be higher than the incremental deployment cost of an integrated SNC/ISC system. Addition of ISC functions to COMS would require an increase in level of redundancy for COMS and addition of interfaces to control nonCCSDS/real-time payload information transfers. Therefore, adding ISC functions to the CDOS/COMS would cost more than adding ISC functions to the SNC system. This leads to the relative cost rankings of 4 and 2 respectively.

Operational cost ranking

The number of SN events per day are projected to increase from approximately 225 per day (in 1990) to approximately 500/day in the year 2000. The current NCC approach for ISC functions uses a dedicated person to monitor each event, while the automated approaches use an operator to process exception alarms only. Therefore, the current manual approach was judged to be the most expensive alternative, i.e. ranked as 4.

The Second TDRSS Ground Terminal (STGT) design uses the concept of continuous monitoring with exception alarms and system reconfiguration (if needed). Adding the ISC exception alarms to this system was judged to result in lowest operational costs due to the incremental nature of this cost element. Therefore, this alternative has been ranked as 1.

Operating a standalone ISC system (either at GSFC or WSC) was judged to incur higher operational costs compared to the incremental cost (for ISC functions) of an integrated CDOS/COMS system. Therefore, the integrated CDOS/COMS alternative has been ranked as 2 and the standalone ISC system as 3.

4.4 Risk

Risk refers to those aspects of the system architecture that, when unmitigated, could seriously impact the ability to implement the system in a timely and cost effective manner. Typically risk areas are characterized as either performance risk, technical risk, schedule risk or cost risk. Performance risk refers to the inability of portions of the system to meet their assigned timing budgets. Technical risk is related to the specification of a system function that simply can not be realized with existing technologies. Cost and schedule risks are related to the inability to develop the system within allotted budget and time constraints.

Because this study is architecture related, performance and technical risks are primarily addressed in the following subsections. Risks are identified in order of importance with respect to the possible impact on SN.

4.4.1 Resource Allocation

4.4.1.1 Demand Access

For demand access there is a performance risk associated with the ability to accurately project user resource request levels, because utilization must be kept low for acceptable service. An underestimation of the ratio of available resources to requests could seriously impact the general utility of a demand access service. Periodic modeling of expected network access based on evolving user needs provides a means of detecting this problem before it affects actual network operation.

Similarly, there is a performance risk associated with the ability of the SNC to meet user needs in a failure condition. Specifically, it is plausible for a resource failure (i.e. loss of a Data Relay Satellite) to render the demand access service ineffectual.

4.4.1.2 Pre-planned

The generation of optimal shift requests is currently a technical risk in the pre-planned resource allocation alternative. This process is currently done manually and the feasibility of automating this function needs further analysis. An option to total automation, however, is to leave the generation of shift requests a manual function with decision aids provided to support the process. The generation of a proof-of-concept SNC planning prototype would address this and possibly other related risks.

The acceptability of a fluid scheduling concept to the SNC user population is a risk because some users may be reluctant to adopt a new approach. The possibility exists that the current 'low priority' users may need freeze points several days in advance while the high priority users may have freeze points only hours in advance. In this case, either the higher priority users could be inappropriately blocked, or the lower priority users pre-empted after making possibly days of preparations (e.g., producing time sensitive command loads). Either result may be unacceptable and additional analysis is needed to resolve this concern.

4.4.1.3 Short-term

The continuous production of short term schedules introduces a possible performance risk. Scheduling is a computationally intensive process and with this alternative, must be done in real-time. The integration of newly arriving user requests with the existing schedule could require a significant level of processing power. The frequency of scheduling activity that would allow the system to be responsive without requiring excessive computational resources would need to be determined, either through prototyping or modeling.

4.4.1.4 Hybrid Access

The hybrid access option suffers from a combination of the risks associated with both the demand access and the short term scheduling alternatives. However, it is less sensitive to the risk of low traffic projections because it is possible to fall back to a purely pre-planned mode.

4.4.2 Resource Partitions

There is a general risk concerning the ability to specify the schedule for a subset of the SN resources as unclassified. Additional discussion and analysis of system security requirements are needed to verify the viability of this approach.

In the case of partitioning, the occurrence of a system failure (i.e. loss of a Data Relay Satellite) would cause a resource shortage. In this situation, the ability of the system to maintain a partitioned service across two subnets could be significantly degraded. The two partitions might need to be merged into a single service. This could have a serious impact on the utilization of the operational resources as well as the percentage of requests that could be satisfied.

There is a possible need for a 'guard' processors at both GSFC and WSC. This is due to the likelihood that user POCCs will be needed to pass user spacecraft interface information directly to WSC in the SNC timeframe. Since WSC will be a classified facility, security functionality (redundant with that in the SNC) will be needed to ensure that classified information is not accidentally returned to the POCCs. The retrofitting of security functionality into WSC could introduce unanticipated cost and schedule impacts into those systems.

4.4.3 Real-time vs. Non-Real Time Functional Partition

The Real-time versus Non-real Time Functional Partition purely represents an allocation of SNC functions across physical subsystems. This represents no obvious risk to the development or operation of the SNC (beyond those associated with resource allocation and computer security identified above) assuming a robust communications mechanism exists between the subsystems.

4.4.4 Integrate SNC with ATGT

A performance and cost risk results from the integration of the SNC and ATGT functionalities. In general, the integration of both new hardware and software becomes more difficult and costly due to the interdependency of the systems. The additional processing load placed on the ATGT processors could seriously limit their planned evolution in the SNC timeframe. A thorough engineering analysis of this option would need to be performed to evaluate its feasibility.

4.4.5 ISC

4.4.5.1 Manual Intersystem Coordination

A general risk with man-in-the-loop systems is the responsiveness the system can achieve. With the number of projected 'events' doubling over the current load in the ATDRSS timeframe, the ability of an operator to generate timely and consistent responses to problems is in question. The operators could very easily become overwhelmed during times of peak activity.

4.4.5.2 Automated Intersystem Coordination

For the stand-alone configuration of the ISC, the need for standard inter-system interfaces to support the system is critical. This is viewed as a risk mainly because of the diversity of organizations that would be involved in the implementation of such standards. The impacts on related systems [e.g., CDOS, Flight Dynamics Facility (FDF), NASCOM II] to support an ISC interface for monitoring the system-of-systems could be non-trivial. Reaching agreement on the necessary standards for the Management Information Bases (MIB) by all organizations that operate those systems could be difficult. Furthermore, the standards activities associated with managing "system-of-systems" is lagging.

One of the primary functions of the ISC is the monitoring of the overall SN. The additional processing imposed by such real-time monitoring could introduce a performance risk. Care must be taken in defining a level of monitoring that provides the operators with useful information while not impacting system performance. This level may be initially determined through a set of system modeling exercises.

4.5 Performance

4.5.1 Resource Allocation

Results from the NASA NPAS model and the CTA Optimized Network Engineering Tools (OPNET) model provide the basis of evaluating Tracking and Data Relay Satellite System (TDRSS) resource allocation among the missions. The primary metric for evaluating scheduled and demand access options is blocking probability, or percentage. Options were evaluated for all scheduled traffic and different resource allocation options for some of the missions using demand access. These results are summarized below. Details of the models and results are provided in Appendix E.

As expected the results from the Network Planning and Analysis (NPAS) model for the scheduled access case show a lower maximum blocking probability than the cases where five missions use demand access. Scheduling the use of the TDRSS resources provides more efficient usage of the resources. The maximum blocking probability for the scheduled access case for the 1998 baseline traffic requirements is 10 percent. For the demand access case the maximum blocking probability determined in the OPNET model is around 30 percent. This three fold increase in blocking assumes that demand access requests cannot wait for resource availability (i.e., if a resource is unavailable when a request is made, the request is blocked). A breakdown of the blocking across the resources shows that all of the blocking occurs on the Single Access (SA) resources for the demand access requests. However, the blocking probability can be reduced to zero for the demand access requestors if they use Multi-access (MA) resources only.

Further evaluation of the demand access case shows that blocking can be reduced to below 10 percent if the demand access requestors have the flexibility of waiting for a resource to become available (i.e., a window period). The range for blocking probability is from zero to 30 percent for window size between the full visibility time that a mission has with the TDRSS constellation and zero, or no waiting for a resource. The blocking probability is at six percent for a window size of 50 percent of the total visibility time.

4.5.2 Resource Partitions

In general performance is degraded when resources are partitioned among users. Partitioning resources results in less efficient usage of resources. However, a partitioned scheme could allow blocking to be better managed. If high priority, preemptive users are partitioned onto a separate set of resources, then the lower priority users may face a higher blocking probability, but they will not have the frustration of having to reschedule after a preemption.

The model of demand access showed that partitioning resources by manned versus unmanned missions results in lower blocking probability than if those two type of missions are mixed together in a subnetwork partition. One scenario evaluated included Freedom and Space Transportation System (STS) in the same subnetwork partition as unmanned missions. This case resulted in higher blocking probability for demand access requestors than the case where these two missions were taken out of the unmanned subnetwork. Both Freedom and STS require full coverage during their TDRSS visibility times from single access resources. When these two sets of demands are moved to another subnetwork partition even though two forward and two return SA resources are reallocated to the other partition, contention on the remaining resources is reduced on the remaining single access resources.

4.6 Ease of Use

4.6.1 Resource Allocation

The relative ease of use of the resource allocation alternatives is shown in the following table. Smaller numbers are more desirable in that they represent a greater ease of use.

Current NCC Approach	2
Demand Access	1
Pre-Planned fixed	2
Pre-Planned fluid	2
Short-Term	3
Hybrid Access	2

The Demand Access alternative is regarded to have the greatest ease of use. This follows from the fact that the user interface would be simpler than that associated with the other alternatives. Requests for service would contain little more than the parameters of the required service. Little negotiation would occur, the service would simply be granted in most cases. As a result, less training of SNC personnel and users would be required. The only case in which ease of use for Demand Access would be poor is if the probability of blocking is high. In such a case users would spend excessive time submitting requests. This is analogous to repeatedly dialing a

long distance service, dialing a number, and finally entering the account number only to get a busy signal.

Four of the alternatives are regarded to have equivalently moderate ease of use. These are the Current NCC, Pre-Planned Fixed, Pre-Planned Fluid, and hybrid access alternatives. Both Pre-Planned alternatives will result in a more complex user interface than demand access, with a corresponding amount of training required. However, this complexity is not thought to be greater than that of the current NCC. This complexity is a natural consequence of the additional steps and features required to submit requests in advance and to negotiate resolution to conflicts; since all requests would be pre-planned, the amount of negotiation would be maximum. The corresponding benefit is that the amount of blocking would be minimum compared to the other alternatives.

Hybrid Access would result in the greatest amount of user interface complexity and training since it would have a combination of all the aspects of the On Demand and Pre-Planned interfaces. However, negotiation would be intermediate since some requests would be serviced by the On-Demand access mode others by the Pre-Planned access mode. Similarly, blocking would be a composite of that inherent to the two modes. Given that complexity is high but negotiation and blocking is intermediate the hybrid access alternative is ranked at the same level as the Pre-Planned alternatives.

The Short-Term alternative provides the lowest ease of use. The user interface complexity and training requirements are somewhat less than that of the Pre-Planned alternatives; the amount of information needed in the request would be reduced. However, the shorter time period limits the types of negotiations which can occur. For this reason a higher level of blocking will be experienced than for the other pre-planned alternatives. This coupled with the need for some POCCs to generate command loads several days in advance could make this alternative very difficult to use effectively.

4.6.2 Resource Partitions

Ease of use will be greatest when the schedule or any part of it is unclassified. It will be much easier for users to find acceptable times since they will have more information to work with. Having part of the schedule unclassified also enables peer-to-peer negotiations to be performed.

The only other sense in which ease of use may be affected is in the case of spillovers. When spillovers occur, it may be more difficult for the alternatives in which the classified and unclassified schedules are separated to tell where the service is coming from or why in some cases service is provided but in other similar cases it is not. In the alternative where the classified schedule is maintained at WSC and the unclassified schedule is maintained at GSFC, there could be an ease of use impact if the request process were not standardized. However, assuming that this alternative would be implemented in a user friendly manner the problem would not exist.

4.6.3 Real-Time vs. Non-Real-Time Functional Partition

Partitioning of functions would simplify the activities of SNC personnel since there would be a clear delineation between the real-time and non-real-time functions. However, partitioning of functions could make the activities of users more complex since the user may have to interact with two system elements.

4.6.4 Integration of SNC with ATGT

Ease of use in this context is not as relevant to the users of the SN as it is to the operators of the SN and the associated elements. From the point of view of the operators leaving the systems separate makes the overall system harder to use since there are two systems to operate. Integrating the functions so that they are resident on same processing system makes the overall system easier to use since there is one system to operate.

4.6.5 ISC System

Manual inter-system coordination is the hardest to use for both users and operators. It places a large burden on the operators of the system and it impinges on user's ability to obtain service.

All forms of automated intersystem coordination remove the major impact on the user since operators are freed to perform other tasks needed to satisfy users' objectives. In addition, Standalone Automated Intersystem Coordination is operationally simplest since the operators are collocated with system and they are focused on coordination rather than trying to support multiple aspects of the host system. A standalone coordination scheme may slightly decrease ease of use for users since they will have an additional functional entity to deal with.

Although integrating the intersystem coordination function with SNC, CDOS, NASCOM, or STGT reduces the number of functional entities the user must deal with, it may decrease the system responsiveness to users. Operators will have multiple jobs to do, increasing the likelihood that they will not respond to user needs as rapidly. In addition, the operators may not be collocated with the system supporting intersystem coordination and hence have less understanding of it or direct control over it.

4.7 Flexibility

4.7.1 Resource Allocation

These alternatives do not have a major impact on the flexibility of the architecture. However, the hybrid access alternative is more flexible in that a change in the mode of operation (towards more demand access or more pre-planned) could be accommodated without a change in the underlying architecture.

4.7.2 Resource Partitions

Partitioning of resources adds flexibility to the system since it provides an additional mode of operation. Even if the type of partitioning ultimately required is different than that initially implemented, it will be easier to satisfy the new requirement if that mode of operation is initially designed into the system.

4.7.3 Real-Time vs. Non-Real-Time Functional Partition

In general, partitioning of the real-time and non-real-time functions will result in greater flexibility. This is primarily due to the fact that partitioning will allow any new non-real-time functions to be implemented in a more simple and straightforward manner. In addition, the allocation of new functions / requirements will be simpler. The only disadvantage of this alternative is that some functions may be difficult to implement because they are not purely real-time or non-real-time.

Similarly, not partitioning the real-time and non-real-time functions will result in lesser flexibility because of the additional complexity of mixing real-time and non-real-time functions.

4.7.4 Integration of SNC with ATGT

Integrating the SNC and ATGT will result in lesser flexibility. Adding functions to the ATGT is regarded to be more difficult than adding functions to the "average" system. Since any new functions or requirements would be constrained by the combination of ATGT and SNC, this would result in even less flexibility. Hence, retaining the SNC and ATGT as separate entities, regardless of whether they are resident at GSFC or WSC, results in greater flexibility.

4.7.5 ISC System

The Manual Inter System Coordination alternative is the least flexible since adding functions or requirements impinges on staffing. The Standalone Automated Intersystem Coordination alternative provides maximum flexibility since multiple processors with reserve capacity are available to handle additional functions or requirements. Integrating the Automated Intersystem Coordination with the SNC, CDOS, NASCOM, or STGT results in minimum

flexibility since the existing systems have limited capacity to handle additional functions or requirements, and since the number of processors available to perform those functions or requirements will be reduced.

4.8 Expandability

Expandability refers to the ability of the architectural alternatives to accommodate increased workload without compromising performance or functionality. The increase in workload may occur due to many reasons, including the following:

- o increase in SN resources/data rates,
- o increase in SN resource usage (e.g. increase in % utilization, increase in number of user systems etc.),
- o increase in number of service requests or events (i.e. more events of shorter duration), and
- o increase in complexity of control functions (e.g. new accounting control, charging mechanisms, allocation rules etc.)

Expandability of an architecture is influenced by many factors, including the following:

- o ability to partition or distribute workload,
- o ability to use special purpose computing systems optimized for specific applications or application type (e.g. real-time systems, parallel computing systems, security guard processors)
- o increase in computational complexity as a function of on increase in workload (e.g. an architecture where computational complexity increases in a linear manner is more expandable than one where it increases exponentially).

The expandability analysis in this section is qualitative only. The architectural alternatives have been ordered according to the qualitative judgement of degree of expandability. Rank order "1" is used for the least expensive (i.e. most easily expandable) alternative. At an architectural level, absolute or relative quantification of percent expandability is not possible.

4.8.1 Resource Allocation

The relative ability to expand for the architectural alternatives is as follows:

on-demand	1	(most easily expandable)
preplanned/hybrid	2	
short-term	3	
current NCC approach	4	(least expandable)

The automated alternatives using distributed data and work management are more expandable than the current NCC approach which is manpower intensive and all processing is centralized. Among the automated alternatives, the computational complexity of the demand access resource allocation approach is a linear function of number of service requests. Therefore, it is the most easily expandable alternative. The computational complexity of the short-term approach increases faster (non-linearly) than the preplanned/hybrid approaches as the number of service requests increase.

4.8.2 Resource partitioning

The partitioned architectures are more easily expandable compared to the unpartitioned architecture due to distribution of workload, provided the partitions are of comparable size. An architecture with 40% & 60% resource partitions is more expandable compared to an architecture with 10% & 90% partitions.

4.8.3 Real-time function partitioning

The architectures with separate real-time and non real-time control systems are more easily expandable compared to the architecture without such partitions. In addition to

distributing the workload, the partitioned architecture facilitates use of systems optimized for real-time and non real-time applications.

4.8.4 Integration of real-time control system and ATGT

Integration of the real-time SNC/ISC systems implies use of common computational components (hardware and software), sharing of operational personnel and operator consoles (e.g. X-terminals connected to different control systems via a LAN). It does not require or prohibit implementation of SNC/ISC functions on the ATGT Automatic Data Processing Equipment (ADPE). The integrated control system will support three types of control functions:

- o real-time ISC,
- o real-time SNC, and
- o all ATGT (SN subsystem) control.

The integration of three different levels of control functions in one system will add significant complexity and therefore impose constraints on expandability. The non integrated architecture will be more easily expandable.

4.8.5 Inter-system Control (ISC)

The relative ability to expand for the architectural alternatives is as follows:

Standalone ISC system	1	(most easily expandable)
Integrated SNC/ISC	2	
Integrated ISC and CDOS/COMS	3	
Current NCC approach	4	(least expandable)

The current NCC approach is least expandable due to lack of automated tools for inter-system control, reliance on multiple two-way voice conversation and post-event fault isolation/correction methodology (rather than during the event).

The standalone ISC architectures is most easily expandable compared to the integrated alternatives which integrate dissimilar control systems resulting in added complexity and constraints on expandability. Also the absolute workload for an integrated control system will increase faster than the workload for separate control systems.

The integrated SNC/ISC architectures was judged to be more easily expandable than the integrated ISC/CDOS control system due to higher degree of similarity between the ISC and SNC functions. Specifically the CDOS control system will not support real-time and non CCSDS payload data. It will support significant LAN traffic among various CDOS processing elements. Both the SNC and ISC systems will support the same space-ground data types. The SNC and ISC systems will control widely distributed SN subsystems [ATGT and Ground Network (GN)] and service providers (SN, CDOS, NASCOM etc.) respectively.

5. Conclusions, Recommendations, and Rationale

In this section a set of recommendations relative to each of the major issues with supporting rationale is presented based on the tradeoff evaluation discussed in the previous section. In summary, the results of this study consist of:

- o recommendations on each key issue with supporting rationale,
- o applicability of other systems and new technology,
- o required infrastructure support such as communications,
- o impact of the primary alternatives on other National Aeronautics and Space Administration (NASA) programs.

These results are summarized in Table 5-1.

Two recommendations that involve several key issues are:

- (1) viewing the Space Network Control (SNC) as an element of within a "system of systems" and
- (2) defining the SNC functionality in terms of the Open Systems Interconnection (OSI) Reference Model Management Framework.

The "system of systems" concept is key because one of the functions of the SNC is the inter-system co-ordination among several complex functions. Also, in order to optimize the systems aspects of the scheduling process, the SNC operations concept must address the end-to-end process relative to the scientist, spacecraft controller, and the SNC operator. The use of the OSI Management Framework will facilitate the use of commercial off the shelf (COTS) products and existing management techniques for the SNC ground based components. For example, the OSI concept of a Management Information Base to define the objects being managed is useful for specifying the reporting and monitoring information flow for inter-system control.

The results and recommendations for each key issue are described in more detail in the following sections.

Table 5-1: Summary of Conclusions and Recommendations

- 1. Adopt a Hybrid Scheduling Approach to Reduce the Operational Complexity of Scheduling**
 - o functionally can be operated like telephone network but do not have SA capacity to meet user demands
 - o fluid schedule (emulate job shop scheduling)
 - o smaller time horizon - (CSOC, Blossom Point)
 - o increased use of demand access for MA
 - o use of standby schedule for periods when shuttle is potentially active
- 2. Incorporate Resource Partitions to Isolate Impact of Users (AT&T)**
 - o minimize manned flight impact
 - o may also allow publication of schedule and use of dedicated ground systems
 - o time of day variations (AT&T)
- 3. Further Automate the Entry, Change, and Conflict Resolution of Schedule Data**
 - o semi-automated generation of shift requests with use of supporting decision aids/large screen displays (CSOC)
 - o use of COTS technologies for distributed data management and distributed work management,
 - o allow for the evolutionary introduction of co-operating expert systems
- 4. Automate the Inter-System Control Function**
 - o introduce ISC Manager and agents
 - o status monitored by exception without dedicated operators
 - o provide "big picture" status on demand
 - o perform automated end-to-end test (ATGT-CDOS-NASCOM-POCC) as a standard part of the pre-pass co-ordination
 - o automated fault isolation
 - o automated configuration parameter validation
- 5. Implement a Real-time / Non-Real time Partition of SNC Functionality**
 - o non-real-time at GSFC
 - o integrate real-time SNC with ATGT
 - o introduce stand-alone automated ISC at GSFC
- 6. Introduce Automated Interface Management**
 - o OSI view of network management
 - o use of COTS components
 - o automated interface management with ASN.1 compilers

5.1. Real-time Versus Pre-planned Resource Allocation

Key Issue 1: WHAT PROCESSING/SCHEDULING IS DONE IN REAL-TIME VERSUS PRE-PLANNED? [WHY CAN'T IT BE LIKE THE TELEPHONE NETWORK?]

The primary result derived from the analysis of this issue is that the Space Network (SN) could be functionally operated in a demand access mode, but the ATRSS resources are inadequate to achieve a sufficiently low blocking probability such that users are satisfied. It is recommended that the SN be operated in a hybrid mode with an increasing level of demand access traffic on the Multi-access (MA) service such that the ongoing scheduling workload can be reduced and unforeseen needs can be readily accommodated.

From a functional point of view, the current Second TDRSS Ground Terminal (STGT) is being designed to provide demand access within five minutes of receipt of a request. Thus, the implementation complexity of a demand access scheme is not a major issue. In fact, from an SN total system point of view it is less complex than pre-planned access. Also, the external systems such as the Department of Defense (DoD) Lead Range or the Jet Propulsion Lab (JPL) Deep Space Network (DSN) may still be operating in a scheduled mode, which would force the SN to also schedule services for these networks. Thus, the main points to be addressed are the performance that can be achieved and interoperability with external systems.

The major distinctions between the SN and the telephone network are the number of users, quantity of resources, and impact of a blocked call. In the telephone network there is a large set of users competing for a large set of resources while in the SN there are a small number of users competing for a small set of SN resources. When a call gets blocked in the telephone network, it usually has a small effect on the user. In most cases, the call is not time critical. Furthermore, the user can usually redial in a few minutes and obtain service because with the large user population, a significant number of calls terminate every minute to free up capacity. In the SN the service is time critical because if not provided the users may miss a real-time science event or spacecraft maneuver. If the user retries, it is less likely that another user has terminated service with the small user population. Even if the user could obtain service by retrying, it may be too late because the command set could not be regenerated to reflect the new time epoch. Thus in this environment, it is necessary that the SN provide a service with a very low blocking probability.

In order to quantify the performance of the SN, a set of simulations were conducted comparing the performance of demand access and pre-planned access with various resource partitions. These results indicate that the blocking probability can be substantially reduced by pre-planning. For example, using an ATRSS constellation with six SA channels, the simulations showed that 30% blocking resulted with demand access while there was only 5% blocking with pre-planned schedule. All of the blocking was on the SA service.

If all of the potential demand access users could be moved to the MA service, then blocking could be reduced to less than 1%. Furthermore, if users could queue for the demand access service, blocking could be further reduced to essentially zero.

The use of a short term schedule (8 hours) can theoretically reduce the blocking significantly, but it does not allow adequate time for the users to perform activities like modifying their command set if the assigned time is not exactly their requested time. Thus, this is not a practical alternative.

5.2 Information Interface

KEY ISSUE 2: WHAT IS THE USER-SYSTEM "INFORMATION INTERFACE?"

The major new concepts resulting from this study relating to the "information interface" are the on-line access to the composite schedule for a subset of the resources, and the use of fluid scheduling. On-line access to the schedule is an especially important concept for the MA service because it may provide users a major incentive to move from the Single Access (SA) service to the enhanced MA service. Fluid scheduling is intended to reduce the resource assignment complexity by not freezing the assignments until necessary.

Although it is envisioned that portions of the SN schedule will remain classified, it is possible that a subset of the schedule may become unclassified by partitioning resources. The resolution of this issue is beyond the scope of this study, but it allows for the introduction of major new concepts. Currently users are quite frustrated upon receipt of a reject notice in response to a schedule request without any explanation. By making the schedule available, a large degree of user frustration can be alleviated. In particular, in our user interviews, some users indicated a very high level of satisfaction with the scheduling of the old ground network whose schedule was published.

The scheduling process with on-line access would operate a centralized schedule under the control of the SNC and resident in an SNC database with access to the relevant users; copies of the schedule could be maintained in local Payload Operations Control Center (POCCs). However, it could be supported with varying levels of information distribution. At a minimum, the database would provide only the times available and blocked; users could identify available times that would meet their needs.

At a second level, the scheduling database could provide the identities of the users who have been allocated service. Thus users could co-ordinate and swap assignments. In a broader context, users could be interconnected and provided access to the complete scheduling and planning database of the other relevant users. To facilitate this approach, a data interchange standard for mission planning and scheduling could be developed. This would involve a data model based on an entity-relationship structure. The resolution of this issue requires further study.

The on-line access to the shuttle schedule may also help other users schedule support time when the shuttle is flying. However, after launch the shuttle schedule is predictable for approximately an 8 hour period. If a class of users who can utilize ATDRSS resources with such short notice can be identified, then service to the users can be improved.

The downside of resource partitioning is the degradation in performance. A set of simulation analyses were performed to determine the impact of resource partitioning. These results show that the major issue is the SA service capacity. However, in order to utilize resource partitioning, it is clear that users should be moved to the MA service to the extent practical.

In order to reduce the complexity of the scheduling it is recommended that a fluid scheduling process be considered. Utilizing multiple freeze points, this concept, emulates the "just in time" scheduling concept used in manufacturing. Since events are not frozen until necessary, the impact of changes is minimized. This will improve the overall efficiency of SNC operation. Use of this concept is not intended to reduce the necessary time for the users to generate their commands, but rather to not make it larger than it has to be. The feasibility of this concept depends on whether:

- o a small number of freeze points can be identified with sufficient volume of traffic associated with each.
- o the users with higher priority have predominantly the later freeze points such that they would often pre-empt the users with earlier freeze points.

These points involve a more detailed analysis.

In interviewing several users of the existing Network Control Center (NCC), several enhancements in the operation were suggested. While more evolutionary in nature than new concepts, these suggestions need to be considered in planning the SNC. These enhancements include supporting modification of schedule entries, modification of sets of schedule entries (shuttle), and attempting to schedule support time in the middle of a specified window rather than at the edges. The latter point is especially important because it will increase the use of window requests and provide the scheduling process with more flexibility in making assignments.

5.3 Control

KEY ISSUE: 3. HOW IS THE SYSTEM CONTROLLED (CENTRALIZED/DISTRIBUTED)?

The primary result derived from the analysis of this issue is that neither centralized nor distributed control is best; rather, a hierarchical control scheme should be introduced. The major new concept included is the formalization of the end-to-end control and co-ordination function currently done in the NCC.

In the analysis of this issue it was recognized that a simple distinction between centralized and distributed control was not adequate. The structure of control elements needed to be formulated and the functions analyzed. Therefore, a control taxonomy was derived (Appendix D) and applied to each of the OSI management functions.

As discussed in Section 2.5, the SNC functions can be categorized as:

- o STGT subsystem functions
- o inter-system control (ISC) functions
- o real-time SNC functions
- o non-real-time SNC functions
- o gateway functions for the interconnection with the international partners.

The hierarchical control structure of these components is analyzed below. If resource partitions are employed with separate SNCs for each partition, control would become distributed. However, the control would be hierarchical within each SNC as described below.

The management of the Advanced TDRSS Ground Terminal (ATGT) will be very complex. Currently the White Sands Complex (WSC) provides reliable service once a user gets access to it. With the introduction of STGT, this will likely improve based on our high level review of STGT. Therefore, this study builds on the control structure proposed for STGT rather than changes it. The STGT subsystem functions should remain resident in ATGT.

In this approach ATGT will have primary responsibility for the major management functions that it is currently performing and will report summary status to the SNC function and ISC function. The non-real-time SNC function would still generate the schedule and perform the high level resource allocation of users to satellites and antennas. This information would be periodically forwarded to the ATGT that would perform the allocation of ground assets to contacts. ATGT would perform fault detection, isolation, and repair of its assets and report status to a higher level for inter-system control. How ATGT performs its fault management and isolation is transparent

to the higher layers, i.e., its operation is "encapsulated." Similarly, the real-time SNC function would perform the high level resource assignment for demand access requests while the ATGT perform the allocation of ground assets.

This inter-system control function would be responsible for the co-ordination, rather actual control, of the SN, Customer Data Operations (CDOS), NASCOM, ATGT, user POCCs, and other external systems. Its focus will be on the interoperation of these systems rather than on their internal operation. The ISC structure will consist of:

- o an ISC manager for the overall co-ordination of the systems and ISC
- o agents resident in each of the systems to perform monitoring, test, and reporting under the direction of the ISC manager.

The major ISC functions are the end-to-end real-time management of faults and performance. In particular, it will be the centralized location where an operator can obtain the "big picture" of the end-to-end real-time system status. Other functions may be added, such as an integrated accounting system, but these are viewed as less crucial to the successful operation of the SN.

The remaining SN control function is the gateway function providing the interconnection of the SN with the international partners. This consists of communications functions equivalent to those specified in the OSI Reference Model, and potentially, a message translation function. This would also be controlled by the ISC.

The automation, allocation, residency, and integration with ATGT of these functions are addressed under key issues 4, 5 and 6.

5.4 Automation

KEY ISSUE 4: WHAT PROCESSING IS AUTOMATED VERSUS MANUAL?

- o **USER SERVICES**
- o **OPERATIONS AND MAINTENANCE**

The major results from the analysis of this issue are that increasing levels of automation can be introduced into the SN for both user scheduling, network operations activity as well as the generation of interface software.

In the scheduling area, the handling of conflicts by voice co-ordination would be largely replaced in an incremental fashion with the automated generation of shift requests, distributed data management to concurrently update schedules, distributed work management to co-ordinate the handling of the shift requests by people, and ultimately by the introduction of co-operating expert systems to execute the shift requests. The order of incremental introduction of these technologies would be defined in terms of increasing risks. With current technology, the user acceptance of distributed work management and the maturity of co-operating expert systems are the biggest risks. Therefore, the initial capability would be to provide the distributed data management of the schedule data and the automated generation of shift requests. This would be followed by the introduction of distributed work management and later by co-operating expert systems.

In the operations area, the concept of an inter-system control center is recommended with the automated monitoring and analysis of the network being performed on an exception basis such that operators are only notified when a problem detected. An operator will be assigned to each contact to ensure that the quality of SN service is maintained, but the operator will be supporting multiple contacts concurrently. In order to do this, the operator must be able to obtain the "big picture" of the real-time system status on demand. The capabilities envisioned to be automated are pre-pass testing by the ISC agents, reporting of summary status by the ISC agents to the ISC

manager, analysis of the status data by the ISC manager, generation of alerts by the ISC manager, and display of the system status by the ISC manager.

The automation of the ISC is based on the OSI concept of a Management Information Base (MIB) for each interface being monitored to define the objects being managed. At the current time, the standardization process has only defined MIBs for components rather than systems. Thus, this is a risk area that needs near term attention because the systems being coordinated are preceding the SNC in their development.

Monitoring by exception is a major change in the operations concept from the existing NCC concept as there will no longer be operators dedicated solely to each pass. Although all of the satellite data collection facilities surveyed had an operator watching each pass, monitoring by exception is achievable with current technology. Since it has not been done, there is a significant risk involved. The first concern is the user acceptance of this approach; locating the ISC operators at Goddard Space Flight Center (GSFC) may facilitate this acceptance. Second, it is imperative that the infrastructure be introduced to support the distributed data management of the ATDRSS configuration so that the SNC and the POCC have the same configuration parameters; operators will not have the time to sort out these parameters under the new concept. Third, the handling of the perturbations introduced by the shuttle will have to be streamlined. This can be done by allowing the users on-line access to a selected subset of the SNC schedule (as discussed in Section 5.3) and using a standby schedule.

It is envisioned that the communications interface software will be largely off-the-shelf OSI based software. This will enable the use of ASN.1 compilers to generate new encoding/decoding software when interfaces are modified. This will make changing interfaces simple and efficient, which could be extremely valuable in developing interfaces for the international partners.

5.5 Residency

KEY ISSUE 5: WHERE IS THE DATA PROCESSING PERFORMED? WHERE IS THE DATA STORED?

The major conclusions of this analysis are that the STGT subsystem functions should remain resident in the ATGT at the WSC and the non-real-time SNC functions should be resident at GSFC in close proximity to the user POCCs. The location of the real-time SNC and ISC functions are dependent upon whether they should be integrated into other systems and are discussed under Key Issue 6 in Section 5.6. The location of the international gateway is not a driving issue and is left TBD.

As discussed in Section 5.4, many of the SNC functions were assigned to the ATGT. The guiding principle in doing this assignment was "encapsulation" such that the low level control functions could be performed near the components being controlled with only higher level status being reported to higher levels. The overall rationale for this approach is that:

- o the complexity of the SN dictates the use of a such an approach
- o existing service availability is excellent
- o management approach for the STGT looks very good and will only improve service.

Thus, there is no rationale to migrate functions from the STGT in the ATGT era or to even consider the specifics of how STGT/ATGT performs them.

The non-real time subsystem, primarily performing the resource allocation function, would be located at GSFC. This is recommended because the schedule processing functions are functionally different (transaction processing vs. real-time

communications) from the STGT functions and are not prime candidates for integration into ATGT. Furthermore, the primary traffic flows would be between the user POCCs at GSFC and the non-real-time subsystem. Although the traffic flows are not large compared to science data, the users will not be affected by congestion or failures in NASCOM when accessing the non-real-time system (if it is located at GSFC). This is especially important with the recommended increase in the level of SNC automation. Therefore, the non-real-time subsystem should be located closer to the users at GSFC.

Within the non-real-time subsystem, the allocation of the generic scheduling system to either the SNC or the POCC is unresolved at this time. As discussed in Section 3, the primary alternatives are either to build a centralized generic scheduler or provide POCCs a set of reusable software modules that can be tailored to their individual needs. The primary issue to be addressed is the complexity of a centralized generic scheduler and requires further analysis.

The location of the real-time SNC has virtually no impact on the communications costs based on the projected data flows. Whether the real-time SNC is located at GSFC or WSC, real-time status information will be brought back to GSFC for delivery to the user POCCs [or to Johnson Space Center (JSC) and other centers]. Furthermore, the reduction in staff that may be achieved by locating the real-time SNC at WSC is small compared to the reductions that can be achieved by implementing monitoring by exception. Whether the intra-SNC and inter-system control be resident at White Sands is largely dependent on the benefits of their integration in ATGT as discussed in Section 5.6.

The location of the gateway(s) for interconnection with the international partners is largely dependent upon the volume and the destination of their data flows. Technically, they can be modularly integrated anywhere. Thus, this issue is left for more detailed analysis based on these traffic flows.

Another factor that may ultimately affect the residency of the SNC functions is whether the SN assets are partitioned. If so, this opens the possibility of partitioning the subsystems for classified and unclassified schedule processing. This is largely a cost issue since major cost reductions can be achieved if SN facilities may be operated as unclassified. The resolution of this issue requires access to classified data and is beyond the scope of this study. As long as part of the schedule can be published the user needs can be satisfied.

5.6 Absorption of SNC by Other Systems

Key Issue 6: CAN THE SNC BE ABSORBED BY OTHER SYSTEMS?

The most attractive alternative is to integrate the real-time SNC functions into the ATGT and introduce an automated ISC at GSFC. Further validation involving an integration complexity analysis of STGT and a timing and sizing analysis is required but is beyond the scope of this study. The major alternatives considered relative to this issue were:

- o integration of the real-time SNC and/or the ISC functions into the ATGT
- o integration of the ISC functions into either CDOS or NASCOM control centers.

The analysis of these alternatives is summarized below.
From a functional point of view, it is natural to extend the real-time control functions into the ATGT because of the similarity to the functions currently performed. The principal functions that could be potentially integrated are the validation of POC commands to modify the ATDRSS channel configuration during a support and th

handling of demand access requests. Since the existing ground terminal already validates commands, this function should be integrated into ATGT. The tradeoffs associated with the integration of the demand access function are more complex and depend on specific functionality. First, a single point of processing in ATGT would have to be established to allocate resources so the user would not have to know which ground terminal to access; this requires an upgrade to the STGT architecture. Second, if demand access is a simple function providing service on a First In First Out (FIFO) basis, then it is reasonable to integrate it into the ATGT. However, this function will be more complex if queueing of demand access requests is performed, and the real-time SNC performs some "look ahead" processing in order to optimally allocate SN resources. In this case, it is less attractive to integrate demand access into ATGT. In general, the integration of SNC components into ATGT will provide limits on the flexibility and expandability of the SNC.

Another major issue to be considered in this analysis is the upgrading of the STGT security functionality. If the real-time SNC functions are integrated into ATGT, then ATGT will communicate directly with unclassified POCCs in a transaction mode. Since STGT does not have this capability, its security architecture will have to be upgraded with the introduction of a Restricted Access Processor. This introduces some risk.

To resolve this issue, a more detailed design analysis of the integration complexity and performance is required. This would involve an analysis of the software hardware integration complexity and the capacity of the ATGT processors and Local Area Networks (LANs).

The issues associated with the integration of the ISC into the ATGT are similar. The STGT is planning to perform a real-time test of its equipment prior to each pass. This could be extended to perform the end-to-end pre-pass test discussed in Section 5.4. However, the risks involve retrofitting security and the integration complexity as discussed above. ATGT will be a classified system and would have to be enhanced to perform co-ordination with unclassified systems.

The other major issue associated with the integration of the ISC into the ATGT concerns transition. As discussed in Section 5.2, it is intended to modify the SNC operations concept to a monitor by exception mode. To facilitate this transition, it would be desirable to have the ISC located at GSFC such that inter-personal communication between ISC and user personnel on either a periodic (e.g., weekly) or emergency basis would be easier.

CDOS is being designed to accommodate only Consultive Committee for International Telegraph and Telephone (CCSDS) missions while the ISC must accommodate non-CCSDS missions as well. On the other hand, NASCOM II will be a general utility supporting applications other than SN, but it will be providing only a basic communications service. Thus major enhancements would be required for integration of the ISC into either CDOS or NASCOM II. Furthermore, the ISC may be a classified system while CDOS and NASCOM II are unclassified systems; it would be undesirable to extend the security requirements into CDOS and NASCOM. Therefore, these approaches are not recommended.

Historically, the NCC has performed the ISC functions manually. Furthermore, the ISC functions could be integrated into the real-time SNC design from the ground up rather than retrofitted into the designs of other systems. Thus, it is attractive to integrate their automation into the real-time SNC providing the latter is not integrated into the ATGT.

5.7 Systems Impact

5.7.1 Infrastructure Support

The major infrastructure capabilities required to support the alternatives recommended above are distributed data management, distributed work management, and OSI communications. These capabilities are described in Section 3.4

5.7.2 Other NASA Systems

Both service users and providers will be affected by the concepts formulated in this study. The primary alternative recommended above affecting other provider systems is the automation of the ISC functionality. The major modifications would be the establishment of a MIB defining inter-system reporting and automation of the ISC agent in these systems for monitoring, test, and reporting.

The user POCCs would be similarly affected by the ISC automation. Also, since ISC would monitor by exception, there would not be routine communication between the SNC and user during a pass. Instead, communication would only be required to handle exception conditions. With this mode of operation, the pre-pass test as well as other monitoring would now be automated.

APPENDIX A

SNC FUNCTIONAL DECOMPOSITION

I. ADMINISTER SN

A. Co-ordinate SN Organizational Interfaces

1. Co-ordinate interfaces between SN subsystems
 - a) Goddard SNC subsystems (if applicable)
 - b) ATGT
 - c) WSC SNC subsystems (if applicable)
 - d) Ground Network Stations (GNI)
 - e) International Message Transfer (SMT)
2. Co-ordinate interfaces with other MO&DSD systems
 - a) CDOS
 - b) Space Data Processing Facility (SDPF)
 - c) NASCOM
 - d) Flight Dynamics Facility (FDF)
3. Co-ordinate interfaces with non MO&DSD service providers
 - a) DoD Lead Ranges (LRI)
 - b) Deep Space Network (DSN)
 - c) ESA (as a service provider)
 - d) NASDA (as a service provider)
4. Co-ordinate interfaces with users
 - a) US POCCs
 - b) ESA
 - c) NASDA
 - d) Other International users providers
5. Maintain Authorized User Database (excluding sensitive/classified information)
 - a) User name, IDs, contact information etc.
 - b) Project SORD (includes negotiated/projected SN usage)
 - c) Configuration Codes for different service types
6. Establish security policies
 - a) Identify and define security events that require logging
 - b) Establish policies for information exchange with external systems
 - c) Access Control policy
 - d) Accountability (audit reporting)
 - e) Assurance policy

B. Provide Technical Operations Direction

1. Co-ordinate SN Systems Engineering & Planning
2. Establish availability objectives
 - a) Identify SN fault events to be managed
 - b) Establish maintenance policies/procedures/schedules etc.

3. Establish performance objectives (QOS and resource utilization)
 - a) Identify SN services and resources for which performance information will be collected
 - b) Distribute performance data reported by SN subsystems
4. Establish resource allocation policies/guidelines for resolving conflicting requests
 - a) Identify SN resources that require explicit allocation/reservation
5. Identify sub-system (ATGT or GNI) resources that will be managed dynamically by the sub-systems themselves.
6. Establish security procedures
 - a) Identify SN resources that require explicit security management
 - b) Classify SN resources/services according to the security policy

C. Manage SN Human Resources

1. Recruit Staff
2. Allocate Staff
3. Administer Personnel
4. Oversee Training and Certification
5. Administer career planning
6. Manage security clearances

D. Perform SN fiscal planning

1. Collect SN costs
 - a) Human Resources costs
 - b) Equipment, facilities and services costs
2. Obtain resource utilization information
 - a) historical data
 - b) future utilization/demand projections
3. Establish chargeback/billing policies, i.e. different rates for different service types
4. Establish accounting management objectives, report formats, reporting frequency (period) etc.
5. Compare actual usage/activity against negotiated/projected usage and take necessary corrective actions, if anomalies found

II. OPERATE SN

A. Provide User Services

1. Receive "scheduled space-ground service event" request from "Manage Resource Allocation", validate request and verify availability of necessary subsystem resources
 - a) SSA Forward Service (SSAF)
 - b) SSA Return Service (SSAR)
 - c) KuSA Forward Service (KSAF)
 - d) KuSA Return Service (KSAR)
 - e) KaSA Forward Service
 - f) KaSA Return Service
 - g) Tracking Service
 - h) Multiple Access Service (MA)
 - i) S-band contingency service
 - j) Ground Network (GN) services
2. Provide SMT services for international partners
3. Pre-event co-ordination
 - a) Confirm/verify "service event parameters" with the requesting user system
 - b) Notify external service providers
 - c) Assign, initialize and activate necessary SN resources
 - d) Perform end-to-end (loopback/path) tests with external systems
 - e) Inform "Manage Resource Allocation" if the tests fail and the scheduled service cannot be provided
4. Provide the acknowledged service
 - a) Activate all necessary SN resources at event start
 - b) Signal end of event to all external systems
 - c) Release ATGT/ATDRSS resources at end of event
 - d) Signal end of event to "Manage Resource Allocation"
 - e) Post-event co-ordination: provide event summary record/information to "Manage Accounting", "Manage Resource Allocation", CDOS and the "User System"

B. Manage Configuration

1. Maintain SN resource allocation Rules Database

2. Maintain Planned Resource Availability Database

- a) resources (taken) out of service
- b) resources already reserved/allocated
- c) resources available for allocation

3. Maintain Preplanned Service Request Database and Scheduled Service Event Database

- a) Receive, validate, log and acknowledge preplanned SN service requests - new/cancellation/change
 - a) External users (US POCCs and Int'l users)
 - b) Internal SN users
 - (1) Simulation and testing
 - (2) System upgrades
 - (3) Maintenance
- b) Translate generic/flexible service requests into specific service events
- c) Negotiate resource availability/allocation among all requestors using the Rules Database, log negotiations record
- d) Schedule service and allocate/reserve SN resources for the service events (ATGT/ATDRSS, GNI, SMT), log and provide event schedule to the requestor
- e) Publish weekly/monthly projected plans for other service providers (CDOS, NASCOM, ESA, NASDA, DSN, DOD LR and FDF)
- f) provide periodic request summary (past, future, performance metric) reports
- g) provide on-demand reports

4. Manage On-demand Service Request

- a) Receive, validate, log and acknowledge on-demand SN service requests
- b) Receive, validate, log and acknowledge service parameter set-up/change requests
- c) provide periodic request summary (past, future, performance metric) reports
- d) provide on-demand reports

5. Manage subsystem Dynamic Configuration

- a) Collect sub-system state information (e.g. in use resources during an event)
- b) Store and maintain **sub-system State Database** (including history)
- c) Control the sub-system state - receive and process requests for state change(s) from "Provide User Services", "Manage Faults", and "Manage Security"
- d) Present the sub-system state

C. Manage Faults

1. Monitor (detect), report/display and record subsystem fault events (Errors, Alarms, Alerts, Anamolies/Unusual trends)
 - a) perform periodic subsystem diagnosucs tests
 - b) obtain information on fault events from lower level module/component/layer management entities
 - c) maintain subsystem **Fault Event Databases**
 - d) maintain external system **Interface MIBs**
 - e) real-time trend displays and alarms
 - f) periodic summary reports incl MTBF, MTTR
 - g) On-demand reports in response to specific requests
2. Analyze subsystem fault event information
 - a) Trace and idenufy faults
 - b) Initiate correction of fault(s) including ATDRSS Interference Resolution
 - c) Report failures that impact QOS to system 'Fault Manager' and 'Dynamic Configuration Manager'
3. Perform layer management (for OSI layers 2 and 3) with external peer systems
4. Manage Inter-system Faults
 - a) Obtain external system Interface MIB information from SN subsystems
 - b) Obtain Interface MIB information from external systems
 - c) Compare, analyze the MIBs for anomalies and other fault indications/trends
 - d) Obtain QOS fault event information from SN subsystems
 - e) Obtain QOS fault event information from external subsystems
 - f) Analyze fault event information and trend indicators and initiate corrective actions
 - a) direct SN subsystems
 - b) co-ordinate corrective actions with external system fault managers

D. Manage Performance

1. Monitor subsystem performance

- a) Collect statistical information under normal conditions
 - a) **Quality of service parameters (QOS)**
 - b) **Resource utilization**
- b) **Record and maintain subsystem Performance Databases**
- c) **Report subsystem performance**
 - a) **real-time trend displays and alarms**
 - b) **periodic summary reports**
 - c) **on-demand historical or special reports**
- d) **Analyze subsystem performance statistics for performance anomalies/bottlenecks and other performance trends/indicators**
 - a) **real-time/short-term trends**
 - b) **Initiate real-time corrective actions as needed**
 - c) **long-term trends**
- e) **Send long term planning input to Tech Ops**

2. Plan and perform subsystem performance tests and simulations

- a) **Submit service requests for allocation of resources**
- b) **Collect QOS and resource utilization information under controlled conditions**
- c) **Record and maintain Performance Test Results Database**
- d) **Analyze aggregate test results and provide planning input to Tech Ops**

3. Monitor end-to-end real-time performance

- a) **Obtain external system Interface MIB information from SN subsystems**
- b) **Obtain Interface MIB information from external systems**
- c) **Analyze the MIBs for performance anomalies/bottlenecks and other performance indications/trends**
 - a) **real-time/short-term trends**
 - b) **long-term trends**
- d) **Direct systems to insert tracer messages and collect transit delay information**
- e) **Analyze transit delay information to predict/isolate transit bottlenecks**
 - a) **real-time/short-term trends**
 - b) **long-term trends**
- f) **Initiate real-time corrective actions, when needed**
 - a) **direct SN subsystems**
 - b) **co-ordinate corrective actions with external system managers**
- g) **Send long term planning input to Tech Ops**

E. Manage Security

1. Manage SN Security

- a) Maintain and manage **Security Administration Database** (including user profiles, key management information, audit criteria etc.)
- b) Provide security administration information to SN subsystems
- c) Collect, record and maintain **Security Event/Audit Database**
- d) Analyze security event information and produce audit reports
- e) Plan and perform SN security tests/simulations

2. Provide subsystem security

- a) Secure classified information/assets
 - a) Access Control (e.g. Guard Processor)
 - b) Encryption
- b) Monitor and Report "security events" to SNC
 - a) Authorized access/usage
 - b) Failed attempts

3. Manage Inter-system Security

- a) Exchange security administration information (such as SDNS credentials) with external systems as applicable/appropriate
- b) End-to-end security tests and simulations
 - a) Plan and co-ordinate tests with external systems
 - b) Submit test service request to "Manage Resource Allocation"
 - c) Obtain results from SN subsystems and external subsystems
 - d) Analyze aggregate test results and publish findings
 - e) Initiate corrective actions

F. Manage Accounting

1. Maintain the **Rate Database** using the chargeback/ billing policies
 - a) pre-planned service rates
 - b) on-demand service rates
 - c) urgent/disruptive service rates
 - d) cancellation/change charges
2. Collect, record and maintain **SN Resource Utilization Database**
 - a) SN subsystem resource utilization (from "Provide User Services")
 - b) Report resource utilization by event(s), by user, by user group etc.
 - a) Periodic reporting
 - b) On-demand reporting in response to specific requests
3. Report end-to-end resource utilization
 - a) Collect, record and maintain External system resource utilization measurements/charges
 - b) Compute and report consolidated charges by event(s), by user, by user group etc.
 - a) Periodic reporting
 - b) On-demand reporting in response to specific requests

III. SUSTAINING ENGINEERING

A. Inter-system Interface Configuration Management

B. Simulate and test SN (for system upgrades etc.)

1. SN Simulation and Tests

- a) Planning
- b) Request allocation of resources to conduct simulation/test activities
- c) Initiate Tests and collect test results
- d) Analyze and publish test results

2. End-to-end simulation and tests

- a) Planning
- b) Request allocation of resources to conduct simulation/test activities
- c) Initiate Tests and collect test results
- d) Analyze and publish test results

C. Maintain SN

1. Maintain SN hardware

- a) Perform periodic preventive maintenance
- b) Perform corrective maintenance

2. Maintain SN software

- a) Identify software problems
- b) Modify/update software

3. Static Configuration Management (such as DOD 483A - hardware/software model/version number, physical location of nodes, types of nodes, applications/functions supported by various nodes, protocol sets supported etc.)

- a) Hardware CM
- b) Software CM

4. Maintain security services/mechanisms (system upgrades etc.)

5. Provide Integrated Logistics Support(ILS)

- a) Manage ILS
- b) Provide SN supply support
- c) Provide SN Technical Data and Documentation
- d) Provide packaging, handling, storage and transportation including security considerations
- e) Provide SN operation and maintenance training
- f) Provide SN support and test equipment
- g) Provide SN facilities support
- h) Secure Physical facilities, equipment etc
 - a) Access Control @ WSC
 - b) Access Control @ Goddard SNC
- i) Provide SN logistics information and computer resources management

D. Provide SN staff training excluding that covered under ILS

E. SN User liason/training



APPENDIX B

Abstract Analogues

One of the preliminary tasks of the SNC study was the execution of an electronic literature search in the area of resource allocation. The goal of this was to identify existing systems and capabilities that addressed scheduling and planning problems that could be readily mapped into the SNC domain. In addition, abstract analogues, based on a mathematical abstraction, were identified and reviewed in an attempt to gain insight into how similar, but non-isomorphic, problems were handled.

Several basic analogues were identified including:

- o the allocation of space to physical items (e.g., components and routes on a printed circuit board);
- o the allocation of physical resources (e.g., tasks to processors in a hard real-time computing environment); and
- o the allocation of time slots (e.g., aircraft landing times).

The initial analysis of the abstract analogues identified several characteristics that could be related to the SNC problem. These characteristics primarily dealt with how some portion of the allocation problem was specifically handled by the analogue. For example, it was noticed that several of the analogues employed multiple allocation policies in order to maximize resource utilization. Specifically, printed circuit board routers use different algorithms to route regular patterns (e.g., memory buses) versus more unstructured connection sequences, and phone calls are routed differently based on the time of day.

However, the specifics of the allocation policies used by the analogues varied slightly with a 'hardest-first' or 'highest priority-first' policy typically in use. This includes giving landing slot preference to airplanes with the longest flight time and scheduling the highest priority task in a computer processing system first. For the most part, these analogous systems also including a method for request priorities to be altered to reflect changes in request status. For example, a flight running low on fuel would be granted a landing slot ahead of others, or a task's priority would be increased as it approached a processing deadline. An alternative to these was the "just in time" scheduling used in job shop environments in an attempt to complete a production item just when it is needed. This is because completion of the item too soon introduces unnecessary storage costs, while completion too late can cost future business. All of these alternatives were considered as part of the analysis of the SNC resource allocation function.

Another characteristic common across the analogues was the use of process sensitive metrics to evaluate the success of the resource allocation process. These metrics fall into two basic classes, percent of requests satisfied and total utilization of resources. Similar metrics are suggested for inclusion in the SNC system to validate the utility of alternate allocation policies and architectures (i.e. resource partitioning).

Results:

The results of the abstract analogue analysis effort, as reflected in Table B-1, shows that the SNC resource allocation problem is somewhat unique. None of the identified analogues exactly fit the characteristics of the SNC resource allocation problem. However, as presented above, the analogues did suggest several innovative concepts to addressing potential SNC problems.

Specific suggestions, and the section of this report that they affected, are as follows:

- 1) Multiple resource allocation policies should be considered in the scheduling of events. The policy being used at any one time may vary based on time of day, pending shuttle launch, or other pertinent decision criteria. Specifically, a mixture of scheduled and demand access policies may be appropriate for the SN. (Section 3.2.6, Hybrid Access Allocation)

- 2) A course to fine granularity in scheduling (i.e. demand leveling) may help resolve scheduling conflicts in a more timely manner and increase resource utilization and scheduling efficiency. (Section 3.2.2, Pre-planned)
- 3) Committing to events at the latest possible time may reduce the number of changes needed to develop a schedule and make the scheduling process more efficient. (Section 3.2.2.2 Fluid Scheduling Functionality)
- 4) The partitioning of resources into sets, based on user needs, may improve scheduling performance with limited impact on resource utilization. (Section 3.2.5, Resource Partitioning)
- 5) A 'variable' request priority scheme, based for example on function to be performed, may allow fairer access to network resources.
- 6) The application of a constraint relaxation method of user event specification (i.e. time window with the request) may reduce the number of schedule changes needed by allowing the automatic shifting of requests to maximize resource utilization. (Section 3.2.2.2 Fluid Scheduling Functionality)
- 7) Stand-by events may be allowed to exist in situations where scheduled time is typically unused (e.g., slips in shuttle launch time). (Section 3.2.2, Pre-planned)
- 8) The identification and use of meaningful evaluation metrics to assess the basic operation and the affect of modifications to the network operations will provide a objective basis of analysis.

These concepts were found to be very useful in formulating alternative SNC resource allocation schemes.

Analogue Characteristic Matrix

	SNC	PCB Router	Mission Sequencing	Telephone Circuits	Operating Systems	Air Traffic Control	Job Shop
Allocation Policy							
Highest Priority	X				X		X
Hardest First		X				X	
FIFO			X	X			
Automation Level							
High	X	X	X	X	X		X
Low						X	
Metrics							
Resource Utilization	X	X		X	X	X	X
Percent of Requests Serviced			X	X	X		
Cost of Operation		X					X

TABLE B-1

Analogue Characteristic Matrix

	SNC	PCB Router	Mission Sequencing	Telephone Circuits	Operating Systems	Air Traffic Control	Job Shop
Resources							
Small Quantities	X	X			X	X	X
Large Quantities			X	X			
Time Horizon							
Pre-planned	X	X	X			X	X
Real-time demand	X			X	X	X	X
Time Constraints							
Fixed time	X						X
Time window		X	X			X	X
Resource Demands							
Multiple	X	X	X	X	X	X	X
Concurrent	X	X	X		X		X
Sequential				X		X	X

TABLE B-1

Appendix C: Summary of Site Visits

In order to identify new concepts and technologies for incorporation into the SNC, the following sites were visited:

- o Air Force Consolidated Space Operations Center (CSOC)
- o Naval Ground Station at Blossom Point
- o AT&T Network Operations Center
- o GTE SpaceNet, COMSAT, and INTELSAT

The mission of the CSOC and Blossom Point facilities is satellite data collection while the other facilities and telecommunications carriers. In addition, to these sites the NASA ground system at White Sands, New Mexico was also visited.

The key concepts derived from these visits are summarized in Table C-1 while the findings for each of the visits are summarized in the individual site visit template. Note all of the satellite common carriers (GTE, INTELSAT, COMSAT) utilize centralized control of their networks. AT&T utilizes hierarchical control with the local sites responsible for detection, isolation, and correction and fault conditions while the central site assists in the detection and is able to reroute traffic around failed sites.

From a functional view, the CSOC and Blossom Point Ground Station were the closest to the NASA SNC. However, their utility was limited in this study because the CSOC is less automated than the NASA and Blossom Point is a small self-contained operation with minimal resource conflict.

Details of the visits to these sites are presented in the templates.

REDUCED PLANNING WINDOW TIME HORIZON (CSOC, GTE, BP)

OPEN BOOKING SYSTEM (GTE, COMSAT)

FLUID SCHEDULING (CSOC)

ALLOCATION BY TIME OF DAY PARTITION (AT&T)

LARGE SCREEN DISPLAY OF SCHEDULE (CSOC)

HIERARCHICAL NETWORK MANAGEMENT (AT&T)

RESOURCE PARTITIONING (AT&T, GTE)

CENTRALIZED CONTROL (ALL WITH CSOC MOVING TOWARDS CENTRALIZATION)

MONITOR BY EXCEPTION - COMMUNICATIONS CARRIERS BUT NOT DoD SATELLITE DATA COLLECTION

COLOR CODED FAULT DIAGNOSTIC TREES (BP)

Table C-1: SUMMARY OF KEY POINTS FROM SITE VISITS

ORGANIZATION: AF CSOC

1. CLASS: SATELLITE DATA COLLECTION

2. RESOURCE

TYPE: 13 EARTH STATIONS AND COMMUNICATIONS LINKS

QUANTITIES: SMALL

DEMANDS: MULTIPLE CONCURRENT

3. TRAFFIC TYPES

HEALTH & SAFETY: YES

PAYLOAD: SOME

4. TRAFFIC VOLUME: 10,000 SUPPORTS PER MONTH WITH 40%
RECEIVED DAY OF EVENT

5. SCHEDULING ENVIRONMENT: PRE-PLANNED BUT MANY CHANGES
RIGHT UP TO TIME OF SATELLITE CONTACT; USE 3 DAY PLANNING
HORIZON AND PUBLISH SCHEDULE EVERY DAY

6. TIME CONSTRAINTS: FIXED TIME AND WINDOW

7. FAULT MANAGEMENT

LOCAL: LIMITED CAPABILITY AT GROUND STATION

REMOTE: 15 MINUTE PRE-PASS TESTS CO-ORDINATED FROM
CONTROL CENTERS AT FALCON OR ONIZUKA AFBs

RECONFIGURATION: GROUND SYSTEMS CONTROLLED REMOTELY

8. SYSTEM OF SYSTEMS: YES, REMOTE ANTENNAS/GROUND STATIONS
AND COMMUNICATIONS FACILITIES

ORGANIZATION: AF CSOC (CONTINUED)

9. COMMENTS:

- o CURRENTLY USE MANUAL SCHEDULING WITH 84 FOOT "BUTCHER BLOCK" PAPER WITH MANUAL CONFLICT RESOLUTION; SCHEDULED IS UNCLASSIFIED WITH MISSIONS IDENTIFIED ONLY BY "IRO" NUMBER
- o USE MANUAL VERSION OF FLUID SCHEDULING BY ALLOCATING "HARDEST FIRST"
- o MIGRATING TO AUTOMATING SCHEDULING PROCESS WITH ASTRO, A TOOL WITH LARGE SCREEN DISPLAY OF SCHEDULE AND AIDS TO SHOW VISIBILITIES, GENERATE INITIAL SCHEDULE, AND IDENTIFY CONFLICTS
- o MANUAL PATCHING AND SWITCHING
- o MOVING FROM A DISTRIBUTED ARCHITECTURE TO A MORE CENTRALIZED ARCHITECTURE. COMPUTING RESOURCES HAVE BEEN MOVED FROM ANTENNA SITES TO CONTROL CENTERS AT FALCON AFB AND ONIZUKA AFB. CONTROL FUNCTIONS ARE PERFORMED AT BOTH ONIZUKA (PRIMARY) AND FALCON BUT WILL ULTIMATELY BECOME OVERALL CONTROL CENTER.
- o ULTIMATELY EXPECT TO IMPLEMENT BASCH, A SCHEDULING SYSTEM TO BE RESIDNET ON IBM 3080 HOSTS, BUT THIS IS NOT EXPECTED UNTIL 1992 AT EARLIEST FOR MINIMAL CAPABILITY.

ORGANIZATION: NAVAL GROUND STATION AT BLOSSOM POINT

1. CLASS: SATELLITE DATA COLLECTION

2. RESOURCE

TYPE: SATELLITE GROUND STATION ANTENNAS AT CONTROL CENTER SITE

QUANTITIES: SMALL- SUPPORT 8 LINKS CONCURRENTLY FOR 14 SATELLITES

DEMANDS: SINGLE

3. TRAFFIC TYPE

HEALTH & SAFETY	YES
PAYLOAD	SOME

4. TRAFFIC VOLUME: LIGHT - 20 MINUTES BETWEEN PASSES

5. SCHEDULING ENVIRONMENT: PRE-PLANNED

6. TIME CONSTRAINTS: FIXED TIME AND WINDOW

7. FAULT MANAGEMENT: ALL RESOURCES FOR CONTROL OF GROUND STATION ARE LOCAL; MONITOR BY EXCEPTION; NO PRE-PASS TEST

8. SYSTEM OF SYSTEMS: NO

9. COMMENTS: 80% OF THE TRAFFIC CORRESPONDS TO A SINGLE MISSION WHICH USES AN AUTOMATED SCHEDULING TOOL. THE INPUTS TO THIS MODEL INCLUDE SPACECRAFT STATE, MISSION, TASKING, AND MODELS OF THE SPACECRAFT (THERMAL, POWER ETC.). SOME MANUAL "DECONFLICTION" IS STILL REQUIRED BUT THIS IS GENERALLY MINIMAL.

ORGANIZATION: AT&T NETWORK OPERATIONS CENTER (NOC)

1 CLASS: DOMESTIC TELECOMMUNICATIONS CARRIER

2. RESOURCE

TYPE: TELEPHONE TRUNKS

QUANTITIES: LARGE

DEMANDS: MULTIPLE CONCURRENT

3. TRAFFIC TYPE

HEALTH & SAFETY NO - separate signaling system

PAYLOAD ALL

4. TRAFFIC VOLUME

130,000,000 CALLS PER DAY ON SWITCHED NETWORK

50% USE 800 SERVICE

**5. SCHEDULING ENVIRONMENT: REAL-TIME WITH PREPLANNED ROUTES
VARIED BY TIME OF DAY PLUS DYNAMIC ROUTES**

6. TIME CONSTRAINTS: NOT APPLICABLE - USER RETRY

7. FAULT MANAGEMENT

LOCAL: RETURNS STATUS REPORTS TO NOC EVERY 5 MINUTES

REMOTE: NOC MONITORS TRAFFIC PATTERNS

**RECONFIGURATION: REROUTING PERFORMED BY CENTRALIZED
NODE BUT LOCAL NODE MANAGES ITS CONFIGURATION**

8. SYSTEM OF SYSTEMS: NO

**9. COMMENT: MAINTAIN A PARTITION OF RESOURCES FOR SWITCHED
NETWORK AND LEASED CIRCUITS. THIS PARTITION IS BEING
LESSENERED WITH THE INTRODUCTION OF THE "SOFTWARE DEFINED
NETWORK."**

ORGANIZATION: COMSAT

1. CLASS: DOMESTIC COMMUNICATIONS CARRIER WITH INTELSAT INTERCONNECTION

2. RESOURCE

TYPE: VOICE, DATA, AND VIDEO CHANNELS

QUANTITIES: MEDIUM

DEMANDS: SINGLE

3. TRAFFIC TYPE

HEALTH & SAFETY: YES FOR COMMUNICATIONS SATELLITE

PAYLOAD: MONITORED ONLY

4. TRAFFIC VOLUME: NOT APPLICABLE

5. SCHEDULING ENVIRONMENT: PRE-PLANNED

5. TIME CONSTRAINTS: FIXED TIME

6. FAULT MANAGEMENT: CENTRALIZED AT WASHINGTON, D.C.
CONTROL CENTER

7. SYSTEM OF SYSTEMS: TO SOME EXTENT WITH INTEROPERATION
WITH NATIONAL CARRIERS

8. COMMENTS:

ORGANIZATION: GTE SPACENET

1. CLASS: COMMERCIAL COMMUNICATIONS CARRIER

2. RESOURCE

TYPE: VOICE AND VIDEO CHANNELS WITH FUTURE DATA SERVICE

QUANTITIES: MEDIUM

DEMANDS: SINGLE

3. TRAFFIC TYPE

HEALTH & SAFETY: YES FOR COMMUNICATIONS SATELLITE

PAYLOAD: MONITORED BUT DELIVERED VIA USER GROUND
STATION

4. TRAFFIC VOLUME

5. SCHEDULING ENVIRONMENT

REAL-TIME: ONLY FOR FUTURE DATA SERVICE

PRE-PLANNED: VOICE CHANNELS LEASED IN ADVANCE, BUT
AUTOMATED BOOKING SYSTEM WITH 3 DAY TIME CLOSING
TIME ALLOCATES VIDEO

6. TIME CONSTRAINTS: FIXED TIME

7. FAULT MANAGEMENT: PERFORMED CENTRALLY AT McLEAN
CONTROL CENTER. PERSONNEL AT REMOTE SITES TAKE ACTION
INDEPENDENTLY ONLY IN EMERGENCY.

8. SYSTEM OF SYSTEMS: NO

9. COMMENTS:

- o VIDEO BOOKING SYSTEM, REFERRED TO AS STRATS, PROVIDES
REMOTE ACCESS TO CUSTOMERS AND IS RESIDENT ON A PC CLASS
MACHINE
- o VIDEO BOOKINGS BECOME FIXED 3 DAYS IN ADVANCE AT WHICH TIME
USERS WILL BE BILLED; HOWEVER, RESERVATIONS MAY BE MADE
MONTHS IN ADVANCE
- o VIDEO SYSTEM PROVIDES INTEGRATED SYSTEM FOR BOOKING,
OPERATION, AND SCHEDULING

ORGANIZATION: INTELSAT

1. CLASS: INTERNATIONAL COMMUNICATIONS CARRIER

2. RESOURCE

TYPE: VOICE, DATA, AND VIDEO CHANNELS

QUANTITIES: MEDIUM

DEMANDS: SINGLE

3. TRAFFIC TYPE

HEALTH & SAFETY: YES FOR COMMUNICATIONS SATELLITE

PAYLOAD: MONITORED ONLY

4. TRAFFIC VOLUME:

5. SCHEDULING ENVIRONMENT: PRE-PLANNED

6. TIME CONSTRAINTS: FIXED TIME

**7. FAULT MANAGEMENT: CENTRALIZED AT WASHINGTON, D.C.
CONTROL CENTER**

**8. SYSTEM OF SYSTEMS: TO SOME EXTENT WITH INTEROPERATION
WITH NATIONAL CARRIERS**

APPENDIX D

Formulation of Alternatives - Supplementary Information

This supplementary material was used in the formulation of alternatives. It consists of a definition of a control taxonomy and a functional analysis. First, the control taxonomy is presented in Section D.1. Then for each of the OSI management functions, an allocation analysis is presented in Section D.2.

D.1 Control Taxonomy

To facilitate the analysis of control alternatives, a control taxonomy was formulated. This taxonomy consists of the following elements:

- o Functions of control
- o Elements of control
- o Data Processing modes
- o Decision-making modes
- o Redundancy of control entities
- o Location of control entities

These elements are described in the following sections.

D.1.1 Functions of Control

The ISO/OSI management framework (ISO 7498-4) categorizes the control functions in the following functional areas:

Configuration Management encompasses monitoring, scheduling and controlling the dynamic state (operational status) of the system and its components. For the NASA SN the allocation of SN resources (i.e. scheduling service events) is a crucial part of CM.

Fault Management encompasses fault detection, isolation and the correction of abnormal system operation that (may) cause the system to not meet its operational objectives.

Performance Management enables the system to perform at or above agreed to service performance levels.

Security Management supports the application of security policies.

Accounting Management encompasses collection, analysis and reporting of service utilization by different users.

The full description of these functions can be found in the ISO standard 7498-4. This categorization has been accepted and adapted by many other national and international standards organizations as well as the computer and communications industry. Specific management functions within these broad categories can be provided by a combination of general purpose mechanisms (shared by several functions) and special purpose mechanisms.

D.1.2 Elements of Control

The primary goal of any control system is to assure that the system behaves as expected or planned. This is generally achieved by using a variety of feedback and control mechanisms. The basic elements of a typical feedback and control mechanisms (Figure D-1) are:

Monitoring - the purpose of monitoring functions is to collect the information about the actual system behavior. It can take a variety of forms, e.g., monitoring the signal to noise ratio in a transmission. This element is also referred to as "Data Collection". The monitoring functions generally use the low level modules to observe the system behavior.

Processing - the raw data collected as a result of monitoring a system is often voluminous and describes the behavior of low level modules/phenomenon. This raw data must be reduced and transformed to determine the aggregate system behavior in terms of high level services expected from the system.

Decision making - the processed data is used to compare the actual behavior to the expected/planned behavior. This comparison provides the basis for determining the need for taking corrective actions, i.e. directing the system. The decision making element is also responsible for accepting service requests from the service users and directing the system to provide the services. This is the most visible element of any control system.

Two distinct approaches to control are "centralized" and "distributed". In a completely centralized approach, all monitored information is recorded (kept) centrally and all decisions are made by a central authority (entity). In most large complex systems, a hybrid approach is more effective. In a hybrid approach the degrees of distribution is different for the processing and decision making/directing elements of control. Some of the possibilities are:

- o distributed recording of information with a central authority for decision making,
- o distributed recording of information and hierarchical decision making,
- o distributed recording of information and distributed decision making without hierarchy, i.e., peer-to-peer decision making by consensus.

The following sections describe the alternative approaches for the processing and decision making elements of control.

D.1.3 Data Processing modes

The data collected by the monitoring elements (generally associated with low level modules) must be reduced and transformed. This can be performed in three different modes:

Decentralized - in this mode the monitored data is processed by the entity/module collecting the data.

Hierarchical - this is a variation of the decentralized mode for large multi-level hierarchical systems (such as the NASA SN). Each level receives the processed data from the lower level entities, may process it further and then forwards the processed data to the next higher level. This approach uses the encapsulation concept at each level and only the information needed by the higher level is passed on to the higher level. The relationship between the levels can be dynamic, i.e., the higher level can request or direct the lower level to provide additional/different information.

Centralized - in this mode all data processing is done by a single central entity (possibly redundant). All data monitoring entities send the raw monitored data for processing to the central entity.

Different variations and combinations of the above three modes are generally referred to as hybrid modes.

D.1.4 Decision-making modes

The decision making elements are responsible for:

- o accepting service requests from the service users and directing the system to provide the services,
- o comparing the actual behavior of the system/subsystem under control to the expected/planned behavior and taking corrective actions as necessary.

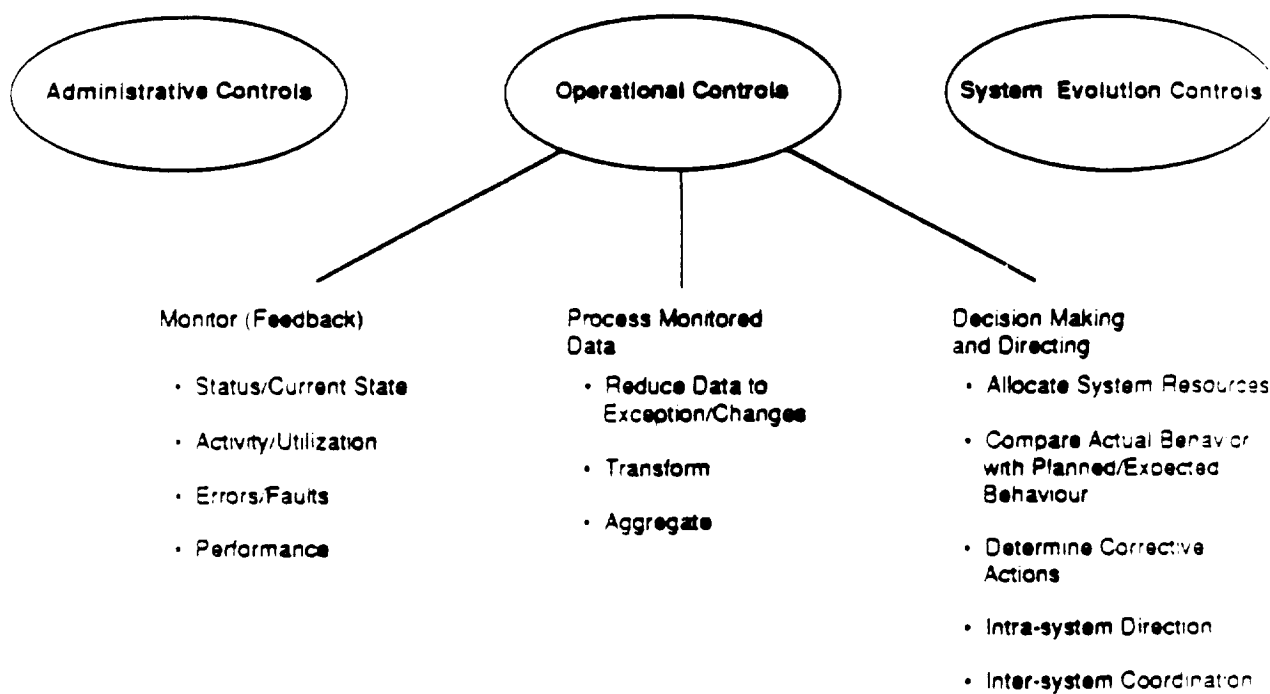


Figure D-1: Elements Of Control

Three commonly used decision making approaches are:

Decentralized peer-to-peer - in this mode there is no central decision making entity. The system is controlled by multiple peer entities (subsystem managers) through a democratic process of consensus decision making. This type of decision making is generally used in a "System of Systems" where the co-operating systems are quasi-independent bounded by policies, mutually agreed upon goals/objectives (e.g. schedules) and QOS (quality of service) parameters.

In theory, peer-to-peer to decentralized decision making can work in conjunction with either centralized or distributed data processing. It is generally used with decentralized processing. Peer-to-peer decision making can use one or more of the following interactions between peers to arrive at a consensus decision:

- o Request-response: a peer sends a request to another peer (or other peers) and expects a response (such as status, request for service granted). The interaction can be repeated to arrive at a negotiated agreement.
- o Periodic messages: a control/management entity can send periodic information messages to other peer entities, e.g. test messages.
- o Exception messages: a control entity may inform its peer regarding the occurrence of unusual/aperiodic events without solicitation, e.g. equipment failure, loss of service or scheduled preventive maintenance downtime etc.

Hierarchical - in this mode there is a clear line of authority structured as a tree. Entities at any level have defined decision-making authority and are responsible for executing the decisions made by higher layer authorities. Each entity reports to the immediate higher level entity and directs the immediate lower level entity. This is the traditional hierarchical organization generally used within a system.

Centralized - this is a special case of hierarchical where there are only two levels. All decisions are made by the single higher level (central) entity.

Different variations and combinations of the above three modes are generally referred to as hybrid modes. For example, the STGT uses peer-to-peer decision making for fault management of redundant equipment, centralized decision making for scheduling of STGT services.

D.1.5 Redundancy of Control Entities

Redundancy of control entities is a crucial aspect in the design of a robust control system for high availability systems, such as the public telephone network. The STGT is an excellent example of a robust system with multi-level redundancy. The current NCC at the GSFC is an example of a redundant control system. Within the context of NASA Space Network Control the issue of redundancy must be addressed at two levels:

- o Intra-SN control entity redundancy and
- o Inter-system control and co-ordination redundancy.

The intra-SN control system redundancy includes the following aspects:

- o Redundancy of centralized control elements, i.e. the control entities that are separate from operational SN subsystems (such as STGT and Ground Network). This is similar to the redundancy of the current NCC at the GSFC.
- o Redundancy of embedded control elements, i.e. the control entities that are embedded in the operational subsystems. For example redundancy of the EXEC ADPE in the STGT.
- o Redundancy of intraSN paths between the control elements.

- o Redundancy of control element locations to protect against catastrophic failures outside the automated system, e.g. earth-quakes and snow storms.

The inter-system control and co-ordination redundancy issues are similar to the intra-SN albeit at the system level instead of at the subsystem level.

D.1.6 Location of Control Entities

For a distributed system or a "System of systems", the location of control entities can have a significant impact on costs, robustness and effectiveness of the control system(s). This is certainly the case for the SNC which will be responsible for controlling SN assets and coordinating SN services with other autonomous systems.

The choice of locations is closely tied to decisions regarding redundancy. Redundant systems can be located in one building (e.g. the existing GSFC NCC), in two different buildings in close proximity (e.g. the STGT and WSGT at WSC) or in two different locations far apart (e.g. WSC and GSFC). Increasing the separation between two redundant systems results in increased robustness at a higher cost. For example, if a severe snowstorm (or fire or earthquake) were to result in a shutdown of GSFC for several days, it will be difficult to operate the redundant system at GSFC.

Some of the alternatives for locating the SNC are:

- o Integrate with ATGT - locate it inside the AGT1 and AGT2 buildings,
- o Resident at GSFC only - no integration with ATGT,
- o Resident at WSC - outside the ATGT buildings,
- o Resident at GSFC and WSC - the WSC SNC systems could be either reside in one of the two ATGT buildings or housed separately.

Several other combinatorial possibilities exist. Another possible approach would be to separate the intraSN and inter-system control systems. In this scenario, the intraSN systems need not be collocated with the inter-system control systems. For example, the intraSN control systems can be housed in the ATGT buildings, while the inter-system control systems can be located at GSFC.

D.2 Functional Allocation

This section documents the rationale for and the recommended allocation of operational control (II. OPERATE SN in Appendix A) functions and associated databases among the SN subsystem control, the intra-SN control (SNC) and the inter-system control (ISC) systems as defined in Section 2.5.

The functions and databases to be allocated were derived earlier as a result of the functional analysis/decomposition as documented in Appendix A (SNC Functional Decomposition). The "Operate SN" function consists of six level 2 functions:

- o Provide User Services
- o Manage Configuration
- o Manage Faults
- o Manage Performance
- o Manage Security
- o Manage Accounting

Sections D.2.1 through D.2.6 address the functional allocation of these six level 2 functions. The functional allocation decisions are generally made at level 3. For example, "Maintain Planned Resource Availability Database" is a level 3 function under the level 2 function "Manage Configuration", which is under the level 1 function "Operate SN". Appendix A also lists lower level functions (level 4 and sometimes level 5) to more clearly define the level 3 functions.

The italicized notation (*SNC , ISC or subsystem functions*) for the level 3 function tables in these sections denote the recommended allocation. The term "subsystem" refers to the three SN subsystems - ATGT, ATDRSS and GN.

The intraSN allocation decisions relates to the degree of autonomy exercised by these three subsystems. The term "SNC system" is used to describe the collection of control entities external to the three subsystems, although they are administratively part of the NASA SN and as such "SNC" will be a SN subsystem. The allocation of control functions described in this section is independent of the location (WSC vs GSFC) and physical implementation details (such as degree of redundancy and integration with ATGT) of the "SNC system".

Key issue #1 (What processing/scheduling is done in real-time versus pre-planned?) is addressed in Section D.2.2 as part of resource allocation functions which are a subset of the "Manage Configuration" function. The other five key issues (#2 through #6) addressed in this study cut across all six functions:

Key issue #2: What processing is automated versus manual?

Key issue #3: What is the User-System "Information interface"?

Key issue #4: How is the system controlled? (centralized/distributed)

Key issue #5: Where is the processing performed?

Key issue #6: Can the SNC be absorbed by other systems?

D.2.1 Provide User Services

The primary purpose of the SN subsystems is to provide services to the users and therefore the majority of functions in this group are allocated to the subsystems. There are three level 3 functions in this group as shown in Table D-1:

II.A.1 Receive and validate "scheduled service event" requests.

II.A.2 Provide SMT services

II.A.3 Pre-event co-ordination

II.A.4 Provide the acknowledged service

Functions II.A.1, .2 and .3 clearly belong to the subsystems. In the current system (pre-STGT era), the pre-event co-ordination function (II.A.2) is performed manually by the NCC operator at GSFC. This process takes about 5 minutes and relies on voice conversation between the NCC operators and the User point of contact.

In the ATGT era CDOS and NASCOM II will be data driven. Thus, the primary focus of the pre-event co-ordination functions will be co-ordination of services between the SN subsystems and the user system for the specific service event. Therefore, this function can be more effectively performed by the SN subsystem providing the service(s) and can be automated to a large degree. The impact of this recommendation on the key issues is as follows:

#2: automate pre-event co-ordination

#4: allocate control of pre-event co-ordination to ATGT

#5: perform pre-event control processing at ATGT

If the pre-event co-ordination identifies a problem that cannot be resolved by the SN subsystem, the resource allocation/scheduling entity should be notified on an exception basis (function II.A.2.e).

Upgrading the Ground Network (GNi) subsystem to perform pre-event co-ordination may not be cost effective in view of the declining use of GNi services. Therefore, the pre-event coordination for GNi subsystem may continue to be handled manually by the SNC personnel or it can be handled manually by the operational staff at the particular Ground Station.

II.A Provide User Services

1. Receive "scheduled space-ground service event" request from "Manage Resource Allocation", validate request and verify availability of necessary subsystem resources (*subsystem functions*)
 - a) SSA Forward Service (SSAF)
 - b) SSA Return Service (SSAR)
 - c) KuSA Forward Service (KSAF)
 - d) KuSA Return Service (KSAR)
 - e) KaSA Forward Service
 - f) KaSA Return Service
 - g) Tracking Service
 - h) Multiple Access Service (MA)
 - i) S-band contingency service
 - j) Ground Network (GN) services
2. Provide Gateway services for international partners (*subsystem functions*)
3. Pre-event co-ordination (*subsystem functions*)
 - a) Confirm/verify "service event parameters" with the requesting user system
 - b) Notify external service providers
 - c) Assign, initialize and activate necessary SN resources
 - d) Perform end-to-end (loopback/path) tests with external systems
 - e) Inform "Manage Resource Allocation" if the tests fail and the scheduled service cannot be provided
4. Provide the acknowledged service (*subsystem functions*)
 - a) Activate all necessary SN resources at event start
 - b) Signal end of event to all external systems
 - c) Release ATGT/ATDRSS resources at end of event
 - d) Signal end of event to "Manage Resource Allocation"
 - e) Post-event co-ordination: provide event summary record/information to "Manage Accounting", "Manage Resource Allocation", CDOS and the "User System"

TABLE D-1

D.2.2 Configuration Management

The most general definition of the term "Configuration Management" includes planning, controlling and monitoring the configuration/arrangement and state of a system and its components. This study does not address the planning functions (part of "Administer SN") and the static configuration management functions (definition/identification of currently approved system as defined in the DOD 483A specifications - these are part of the "Sustaining Engineering").

The operational configuration control functions consist of - resource allocation and dynamic configuration management functions. There are five level 3 functions in this group as shown in Table D-2:

II.B.1 Maintain SN resource allocation Rules Database

II.B.2 Maintain Planned Resource Availability Database

II.B.3 Maintain Preplanned Service Request Database and Scheduled Service Event Database

II.B.4 Manage On-demand Service Request

II.B.5 Manage subsystem Dynamic Configuration

For the NASA SN, resource allocation management is a very critical and complex function. The first four level 3 functions pertain to resource allocation management.

The resource allocation management function in the ATGT era will have a key difference compared to the preSTGT era, i.e. today. The current NCC is responsible for scheduling use of SN, NASCOM, SDPF resources. In the ATGT era, NASCOM would have evolved to the data driven NASCOM II and SDPF functions would be taken over (for the most part) by the data driven CDOS. By definition (of a data driven system) these systems will not require explicit scheduling of services. Therefore, the resource allocation functions will be primarily responsible for allocation of the SN resources only.

There will continue to be a need for coordination of end-to-end service scheduling. These are planned to be implemented by using the advisory functions "Publish weekly/monthly projected plans for other service providers (Function II.B.3.e). This function has been allocated to the SNC (rather than ISC), as part of SN resource management function. The use of the projected plans published by the SNC by other service providers will be optional. In this scenario, the external systems are responsible for keeping track of their resource availability or lack thereof (e.g. due to scheduled downtime or a major failure). This function is in addition to the end-to-end tests performed by the SN subsystems, as part of pre-event co-ordination (Function II.A.3.d) described in Section D.2.1.

In summary, all operational configuration management functions listed in Table D-2, have been allocated to either the SNC or the SN subsystems. No ISC functions were identified.

II.B. Manage Configuration

- 1. Maintain SN resource allocation Rules Database (*SNC functions*)**
- 2. Maintain Planned Resource Availability Database (*SNC functions*)**
 - a) resources (taken) out of service
 - b) resources already reserved/allocated
 - c) resources available for allocation
- 3. Maintain Preplanned Service Request Database and Scheduled Service Event Database (*SNC functions*)**
 - a) Receive, validate, log and acknowledge preplanned SN service requests - new/cancellation/change
 - a) External users (US POCCs and Intl users)
 - b) Internal SN users
 - (1) Simulation and testing
 - (2) System upgrades
 - (3) Maintenance
 - b) Translate generic/flexible service requests into specific service events
 - c) Negotiate resource availability/allocation among all requestors using the Rules Database, log negotiations record
 - d) Schedule service and allocate/reserve SN resources for the service events (ATGT/ATDRSS, GNI, SMT), log and provide event schedule to the requestor
 - e) Publish weekly/monthly projected plans for other service providers (CDOS, NASCOM, ESA, NASDA, DSN, DOD LRs and FDF)
 - f) provide periodic request summary (past, future, performance metric) reports
 - g) provide on-demand reports
- 4. Manage On-demand Service Request (*SNC functions*)**
 - a) Receive, validate, log and acknowledge on-demand SN service requests
 - b) Receive, validate, log and acknowledge service parameter set-up/change requests
 - c) Schedule (or deny) on-demand service
 - d) provide reports (performance/usage etc.)
- 5. Manage subsystem Dynamic Configuration (*subsystem functions*)**
 - a) Collect sub-system state information (e.g. in use resources during an event)
 - b) Store and maintain sub-system State Database (including history)
 - c) Control the sub-system state - receive and process requests for state change(s) from "Provide User Services", "Manage Faults", and "Manage Security"
 - d) Present the sub-system state

TABLE D-2

D.2.2.1 Manage Resource Allocation

Key issues #1 ("What processing/scheduling is done in real-time versus pre-planned?") and #2 ("What is the user-system "information interface") are addressed in Section 3.2. The exact scope of the resource management functions will be affected by the decisions made with respect to the following subjects addressed in Sections 3.2 and 3.4:

- A. How much resource allocation must be done by the SN resource manager and what can be assigned to the SN subsystems? The resource allocation functions assigned to the SN subsystem will become part of the real-time function "Provide User Services". (see Section 3.2.2.4 - Information Interface)
- B. What part of generic scheduling is performed by the SN resource manager and what functions can be assigned to the user system? (see Section 3.4.2 - "Distributed Data Management")
- C. Of the functions assigned to the SN resource manager, what processing/scheduling is done in real-time versus pre-planned? (see Section 3.2)

Item A affects the boundary between the "manage resource allocation" and "provide user service" functions. For example, should the resource manager allocate/reserve a channel (within a group of channels, i.e. a subnet) and let the SN subsystem service provider (such as ATGT) allocate the specific channel. It affects the nature of information exchanged between the two functions. In the example, if the resource manager does not allocate specific channels, then it does need status information regarding specific channels, it only needs information regarding the total number of operating channels.

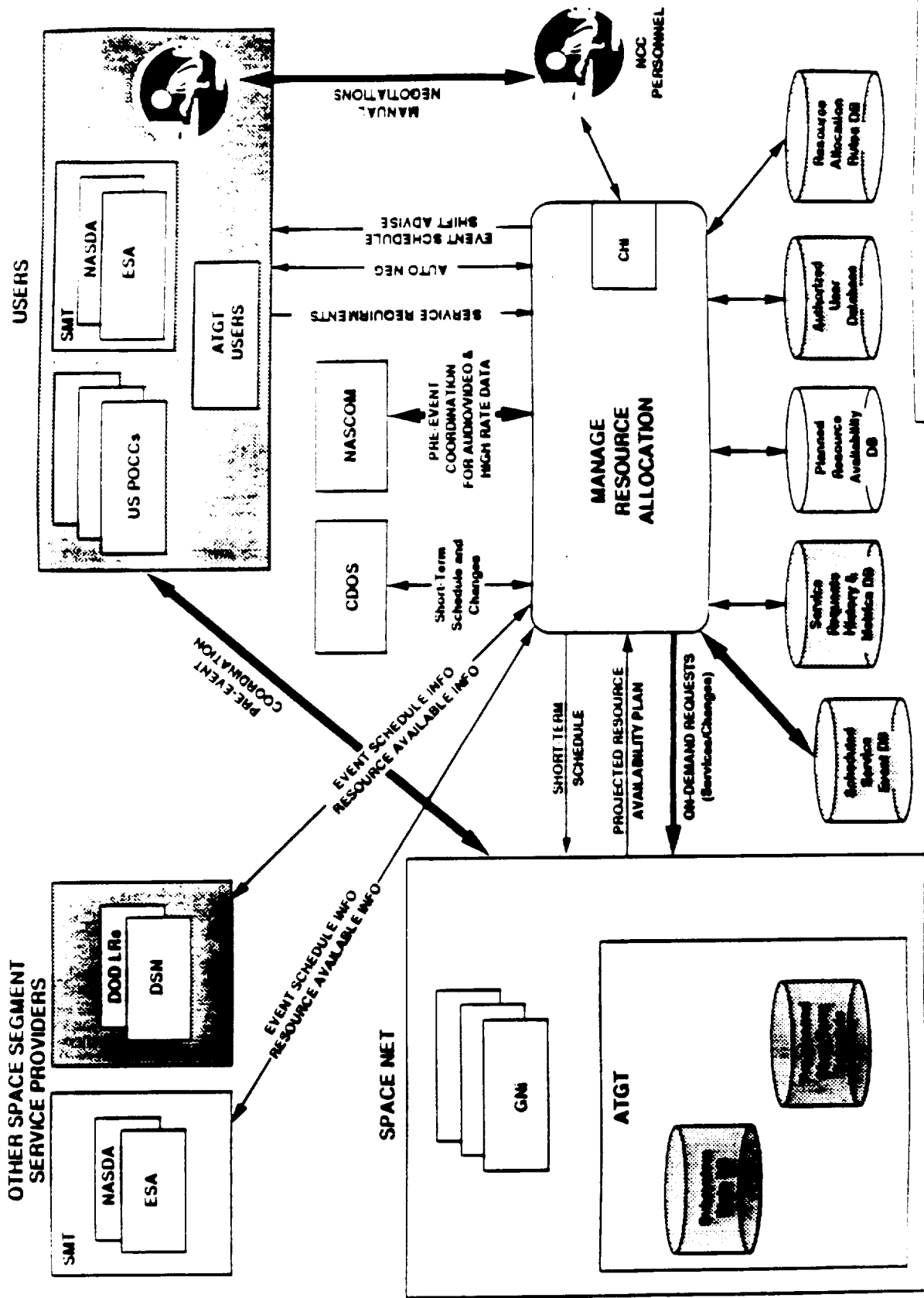
Item B affects the scope and complexity of the algorithms in the functions Maintain Allocation Rules Database (II.B.1) and Schedule SN services (II.B.3.d and II.B. 4.c).

Item C affects the complexity of the resource allocation functions as shown in Figures D-2 and D-3. These figures show the databases and control data flows for the pre-planned/hybrid and on-demand access alternatives respectively.

As shown in the two figures, the on-demand access alternative is simpler than the pre-planned/hybrid alternative for the following reasons:

- o The SNC need not maintain the following databases for the on-demand approach:
 - Preplanned short-term (next n hours) schedule in ATGT, since all service events are scheduled (or denied) immediately after the receipt of the service request,
 - Projected Resource Availability Database, since all service events can be scheduled (or denied) based on the current system state database,
- o The "scheduled service event" database is replaced with the "active event (in progress)" database in case of the on-demand access approach. This is a smaller and simpler database and results into simpler allocation rules for the on-demand approach.

PRE-PLANNED OR HYBRID ACCESS ALTERNATIVE



ON-DEMAND ACCESS

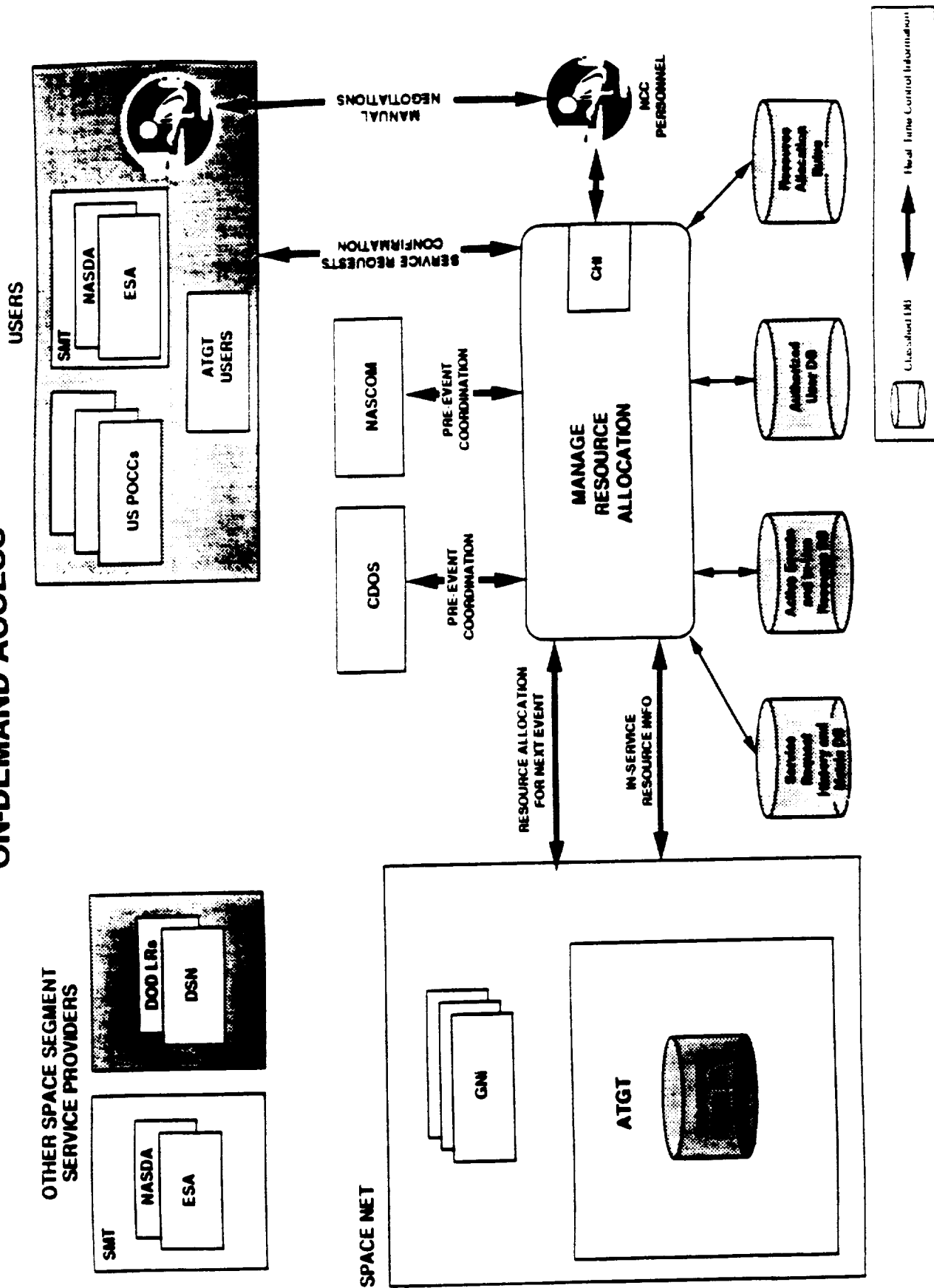


Figure D 3

D.2.2.2 SN Subsystem Configuration Management

The last level 3 function under "Manage Configuration" deals with the dynamic configuration of the subsystems. The dynamic state of a system is affected by - ongoing requests for service (as scheduled by the resource allocation function) and internal changes initiated by fault/performance/security management functions. The dynamic configuration function includes the following level 4 functions:

- o collecting information about the current condition of the subsystems (generally on-demand),
- o obtaining announcements of significant changes in the condition of the subsystems,
- o maintaining the subsystem state database,
- o controlling the subsystem state, i.e. changing the configuration of the subsystem (e.g. taking out a resource for testing and putting it back in service afterwards),
- o presenting the subsystem state (e.g. to the resource manager).

The SNC system is primarily concerned with the overall operational capability of the SN assets, rather than detailed control of low level modules which is best performed within the subsystems (such as ATGT). The dynamic configuration management function specified in the STGT design document meets this criteria quite well and we did not find any reason to change those approach. Therefore these functions have been allocated to the subsystems themselves.

D.2.3 Fault Management

The purpose of fault management is to detect, isolate and correct faults. The term "fault" is used here in a broad sense and refers to any abnormal/anamolous behavior that may impact the quality of service provided by the system. This can includes failures resulting in total loss of service, excessive errors and performance bottlenecks. The definition of fault events is part of the administrative functions under "Provide Technical Operations Direction". Fault management has the following characteristics:

- o It requires continuous monitoring for detection,
- o It can be effectively performed at or above the level at which a fault occurs
- o Performing fault management at higher level results in significant control information flow between components/modules/subsystems which in turn can slow down fault isolation and correction
- o If a fault is managed at the level of occurrence, the only information needed by the higher level (typically system level fault manager) is - impact on QOS (if any) and summary event description (cause, timing, corrective action taken etc.)

The above characteristics imply that an effective fault management strategy would

- o give maximum autonomy to SN subsystems for fault management
- o establish QOS thresholds for the subsystems (e.g. # of lost packets/frames, channel bit error rate, average/maximum delay, availability objectives)
- o use the layer management approach (rather than the system management approach) at OSI layers 2 and 3 for inter-system fault management (e.g. CCSDS SLAP, FDDI SMT)
- o use system fault management for higher layers (OSI layers 4-7) and faults that cannot be managed by the subsystems or layer management.

This approach was used in decomposing the level 2 function "Manage Faults". As shown in Table D-3, there are four level 3 functions in this group:

II.C.1 Monitor, report and record subsystem fault events

- II.C.2 Analyze subsystem fault events
- II.C.3 Perform layer management (for OSI layers 2 and 3)
- II.C.4 Manage Inter-system Faults

The first two (II.C.1 and .2) deal with intraSN fault management, while the last two (II.C.3 and .4) deal with inter-system fault management.

Functions II.C.1 through II.C.3 have been allocated to the subsystems. This is consistent with the STGT design approach, where the STGT subsystem contains extensive fault management capabilities and provides equipment availability messages (SLR) and active service reports (ODM) to the NCC.

The concept of layer management performed by the subsystems (such as ATGT) for inter-system control/co-ordination is a recent development in the communications industry and is generally part of the recent communications standards, such as FDDI SMT. Layer management may not be supported by some of the communication protocols to be used by SN subsystems.

The functional decomposition shown in Table D-3, does not allocate any functions to the SNC function group. It implies that all FM functions are either performed by the SN subsystems or the ISC system. This assumption was made to simplify the analysis and focus the study on answering the six key issues. In a real system, the SNC system will provide a single point of contact for the ISC and other external system, i.e., it will act as the SN manager and communicate with the proxy agent for the ISC manager. Figure D-4 shows the relationship between SN subsystem fault managers and the SN system manager.

The impact of these recommendations on the key issues is as follows:

- #2: uses layer management to increase inter-system FM automation
- #4: gives maximum autonomy to SN subsystems for fault management
- #5: many FM functions are performed by ATGT

II.C Manage Faults

1. Monitor (detect), report/display and record subsystem fault events (Errors, Alarms, Alerts, Anamolies/Unusual trends) (*subsystem functions*)
 - a) perform periodic subsystem diagnosucs tests
 - b) obtain information on fault events from lower level module/component/layer management entities
 - c) maintain subsystem **Fault Event Databases**
 - d) maintain external system **Interface MIBs**
 - e) real-time trend displays and alarms
 - f) periodic summary reports incl MTBF, MTTR
 - g) On-demand reports in response to specific requests
2. Analyze subsystem fault event information (*subsystem functions*)
 - a) Trace and identify faults
 - b) Initiate correction of fault(s) including ATDRSS Interference Resolution
 - c) Report failures that impact QOS to system "Fault Manager" and "Dynamic Configuration Manager"
3. Perform layer management (for OSI layers 2 and 3) with external peer systems (*subsystem functions*)
4. Manage Inter-system Faults (*ISC functions*)
 - a) Obtain external system Interface MIB information from SN subsystems
 - b) Obtain Interface MIB information from external systems
 - c) Compare, analyze the MIBs for anomalies and other fault indications/trends
 - d) Obtain QOS fault event information from SN subsystems
 - e) Obtain QOS fault event information from external subsystems
 - f) Analyze fault event information and trend indicators and initiate corrective actions
 - (1) direct SN subsystems
 - (2) co-ordinate corrective acutions with external system fault managers

TABLE D-3

ORIGINAL PAGE IS
OF POOR QUALITY

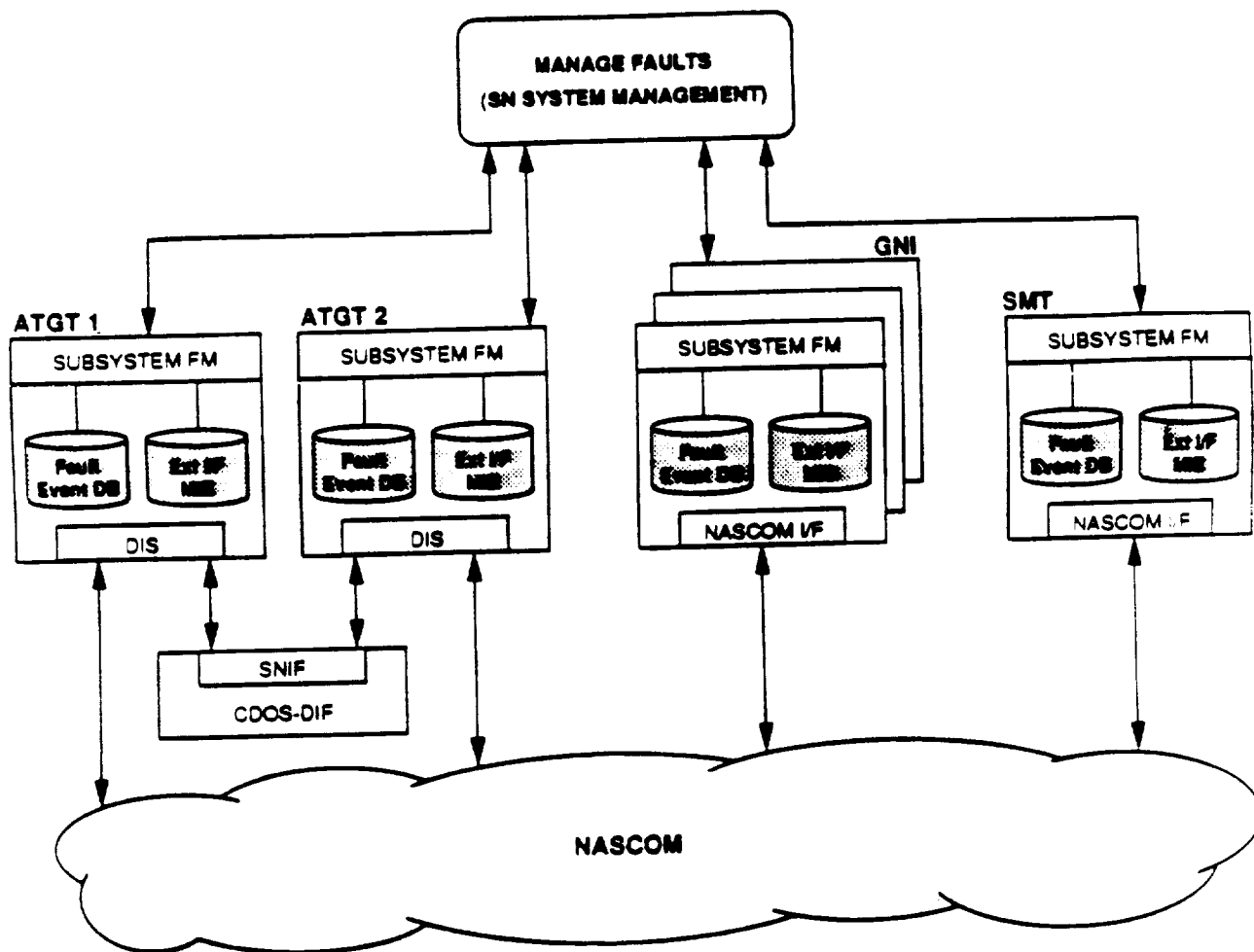


Figure D-4: Intra-SN Fault Management

D.2.4 Performance Management

Performance management functions evaluate the effectiveness of the system as a provider of services and maintain system performance at a specified (agreed to) level. It includes of collection and analysis of statistical information to achieve these objectives. In most systems, it also includes determination of system changes/upgrades to enhance performance. Performance objectives are established by the administrative functions as part of 'Provide Technical Operations Direction'.

There is some overlap between fault management and performance management functions. For example, flow control and buffer management functions can be treated as part of fault or performance management. The results of this study are independent of such definition ambiguities, since the functional decomposition and allocation does not prohibit the use of general purpose management mechanisms common to several functional areas.

Performance management has the following characteristics:

- o It requires continuous monitoring to gather statistical information.
- o It requires infrequent arbitration among lower level modules to resolve conflicts.
- o The statistical information collected is generally massive.
- o Only processed historical data (summary, exceptions etc.) is generally stored and the degree of reduction is increased as the data gets older.

The above characteristics imply that an effective performance management strategy would give maximum autonomy to SN subsystems, as is the case with the STGT design approach. Table D-4 shows the level 3 functions under 'Manage Performance'. The first two address the SN subsystem performance and have allocated to the subsystem performance management entities. These are:

II.D.1 Monitor subsystem performance

II.D.2 Plan and perform subsystem performance tests and simulations

The last level 3 functions (II.D.3 Monitor end-to-end performance) addresses inter-system performance management and is not discussed here.

The proposed functional allocation does not assign any intraSN performance management functions to a system level management entity. A possible exception to this recommendation would be for the data transferred to international partners via the SMT. It is recommended that this determination be made as part of the SMT requirements analysis and architecture development process. The impact of this recommendation on the key issues is as follows:

- #2: the level of automation specified for STGT in the area of PM is adequate
- #3: decentralize performance management, i.e., let each SN subsystem manage its own performance within established objectives. This includes maintaining performance databases.

II.D. Manage Performance

1. Monitor subsystem performance (*subsystem functions*)
 - a) Collect statistical information under normal conditions
 - a) Quality of service parameters (QOS)
 - b) Resource utilization
 - b) Record and maintain subsystem **Performance Databases**
 - c) Report subsystem performance
 - a) real-time trend displays and alarms
 - b) periodic summary reports
 - c) on-demand historical or special reports
 - d) Analyze subsystem performance statistics for performance anomalies/bottlenecks and other performance trends/indicators
 - a) real-time/short-term trends
 - b) Initiate real-time corrective actions as needed
 - c) long-term trends
 - e) Send long term planning input to Tech Ops
2. Plan and perform subsystem performance tests and simulations (*subsystem functions*)
 - a) Submit service requests for allocation of resources
 - b) Collect QOS and resource utilization information under controlled conditions
 - c) Record and maintain **Performance Test Results Database**
 - d) Analyze aggregate test results and provide planning input to Tech Ops
3. Monitor end-to-end real-time performance (*ISC functions*)
 - a) Obtain external system Interface MIB information from SN subsystems
 - b) Obtain Interface MIB information from external systems
 - c) Analyze the MIBs for performance anomalies/bottlenecks and other performance indications/trends
 - a) real-time/short-term trends
 - b) long-term trends
 - d) Direct systems to insert tracer messages and collect transit delay information
 - e) Analyze transit delay information to predict/isolate transit bottlenecks
 - a) real-time/short-term trends
 - b) long-term trends
 - f) Initiate real-time corrective actions, when needed
 - a) direct SN subsystems
 - b) co-ordinate corrective actions with external system managers
 - g) Send long term planning input to Tech Ops

TABLE D-4

D.2.5 Security Management

Security management includes the operational procedures, controls, and system functions required to reduce the risk of unauthorized dissemination of classified materials. These areas of interests are typically segregated into personnel security, physical security, emanation security, computer security, and communications security. This study addresses the computer and communications security only.

Computer and communications security management has the following characteristics:

- o continuous monitoring of all access attempts,
- o need to maintain an audit trail of all access attempts,
- o periodic testing and verification of security mechanisms.

These characteristics imply that the monitoring and audit reporting functions can be performed more effectively by the subsystems. Table D-5, shows the decomposition of "Manage Security" into level 3 and 4 functions. The intra-SN level 3 security functions are:

II.E.1 Manage SN Security

II.E.2 Provide SN subsystem security

The "Manage SN Security" function (II.E.1) includes maintenance of systemwide security administration and audit databases. It also includes analysis of the audit data and testing/verification of systemwide security mechanism. Therefore, this function has been assigned to the SNC. The subsystems are responsible for providing the security mechanisms (II.E.2), as directed by the SNC to assures secure operation of the SN.

II.E Manage Security

1. Manage SN Security (*SNC functions*)

- a) **Maintain and manage Security Administration Database** (including user profiles, key management information, audit criteria etc.)
- b) **Provide security administration information to SN subsystems**
- c) **Collect, record and maintain Security Event/Audit Database**
- d) **Analyze security event information and produce audit reports**
- e) **Plan and perform SN security tests/simulations**

2. Provide subsystem security (*subsystem functions*)

- a) **Secure classified information/assets**
 - (1) Access Control (e.g. Guard Processor)
 - (2) Encryption
- b) **Monitor and Report "security events" to SNC**
 - (1) Authorized access/usage
 - (2) Failed attempts

3. Manage Inter-system Security (*ISC functions*)

- a) **Exchange security administration information (such as SDNS credentials) with external systems as applicable/appropriate**
- b) **End-to-end security tests and simulations**
 - (1) Plan and co-ordinate tests with external systems
 - (2) Submit test service request to "Manage Resource Allocation"
 - (3) Obtain results from SN subsystems and external subsystems
 - (4) Analyze aggregate test results and publish findings
 - (5) Initiate corrective actions

TABLE D-5

D.2.6 Accounting Management

The purpose of accounting management (AM) is to establish charges for the use of SN resources by different users based on their use of the SN services. The most general definition of accounting includes setting limits on use of services by specific users. Examples of limits include maximum usage limits and time of day use restrictions. This study assumes that most limit functions (such as time of day limits) are performed by the resource allocation manager using the allocation "Rules Database". Cumulative usage based limits are handled as part of the administrative functions. The AM functions include the necessary reporting functions. In other words, maximum usage limits are not imposed in real-time, as is the case with many commercial computer services.

At this time the use of accounting management to charge users based on each service event is limited to commercial users (non Government), such as LANDSAT. Use of accounting management is expected to increase in the future as commercial usage increases. The functional decomposition and allocations made under this study are not affected by whether the accounting management is applied to all or selected users.

There is some overlap between manage accounting and manage performance functions as both involve measurement of resource utilization. Accounting controls user or service event specific resource utilization, while performance controls overall resource utilization independent of the users. However, accounting information often includes the level and quality of service provided to specific users for specific service events.

Table D-6 shows the decomposition "Manage Accounting" function into level 3 and 4 functions. The characteristics of accounting management functions are:

- o it is a highly automatable function with minimal manual intervention/assist necessary
- o close interaction with several administrative functions
- o accounting data is sensitive information and requires adequate privacy/security measures
- o accounting data is historical in nature and not necessary for the operation of SN subsystems

The above characteristics suggest that all accounting functions are best allocated to the SNC system.

The impact of the recommended allocations on the key issues is as follows:

- #2: accounting is a highly automatable function.
- #4: centralized accounting.
- #5: perform accounting management at SNC.
- #6: the use of accounting management (a SNC function) is expected to increase in future

II.F Manage Accounting

1. Maintain the Rate Database using the chargeback/ billing policies (*SNC functions*)

- a) pre-planned service rates
- b) on-demand service rates
- c) urgent/disruptive service rates
- d) cancellation/change charges

2. Collect, record and maintain SN Resource Utilization Database (*SNC functions*)

- a) SN subsystem resource utilization (from "Provide User Services")
- b) Report resource utilization by event(s), by user, by user group etc.
 - (1) Periodic reporting
 - (2) On-demand reporting in response to specific requests

3. Report end-to-end resource utilization (*ISC functions*)

- a) Collect, record and maintain External system resource utilization measurements/charges
- b) Compute and report consolidated charges by event(s), by user, by user group etc.
 - (1) Periodic reporting
 - (2) On-demand reporting in response to specific requests

TABLE D-6

APPENDIX E

Performance Model

E.1 Model Description

A simulation model was developed to examine different alternatives of partitioning ATDRSS resources among missions, allocating single access and multi-access resource among resources, and allowing missions to make requests on a demand access basis instead of scheduled. The model was developed using the commercial tool OPNET (OPTimized Network Engineering Tools) and data from the NASA tool NPAS. The model explicitly represents the requests of the group of missions that use demand access, their orbits over a seven day period, and the start of visibility (SOV) and end of visibility (EOV) periods for the different single access (SA) and multi-access (MA) resources for ATDRSS. The orbital and visibilities data are determined from NPAS output data. The scheduled missions' usage of the various SA and MA ATDRSS resources are represented in the model by the busy periods for each resource determined from the output data of NPAS. The performance metrics produced by the model include percentage of requests blocked, percentage blocked on first attempt, average waiting time, percentage waiting, proximity to the middle of the window, and resource utilization. The assumptions of the model are given below.

Two subnetwork partitions are defined for missions: Subnetwork A and Subnet B. Different alternatives are subsequently defined for these two subnetworks. We assume that Subnetwork A missions use only SA resources that are dedicated to them. The remainder of the missions, both scheduled and demand access, use only the remaining SA or MA resources. This model gives results for Subnetwork B users only. The model considers two kinds of demand access requests: immediate, or blocking, and window requests. The former category consists of requests that are lost if they cannot be satisfied immediately. Window requests have a time window that they can wait before they are lost.

The model assumes that the mission traffic load is defined by the third quarter 1998 traffic model developed by NASA and used in the NPAS model. Based on input from NASA modeling personnel the five missions are considered for demand access instead of scheduled access. This mission traffic model is summarized in Table E-1. The NPAS model was executed with the demand access missions taken out. The scheduled request intervals taken from the NPAS model output are assumed to be constraints in allowing the demand access requests to be serviced. Spacecraft masking of antennae during its visibility interval is represented for the scheduled requests but not for the demand access requests. For the demand access requests the model assumes that requests occur only within the respective interval of visibility of the ATDRSS constellation, and that requests occur randomly within that interval but no later than the time after which the request could not be satisfied due to not enough time remaining.

The model assumes that the following SA and MA resources of the ATDRSS era are available: 8 SAF; 8 SAR; 8 MAF; and 20 MAR. It is also assumed that K-band and S-band single access antennae cannot operate simultaneously. A normal operating scenario is assumed (i.e., no routine down times or spacecraft emergencies are represented).

E.2 Evaluation Scenarios

The model was used to evaluate two network partitions. Figure E-1 shows the different parameter categories that were varied in the evaluations. The two partitions were defined by the number of SA resources that are assumed to be dedicated to Subnetwork B missions. The left side of the tree diagram defines model runs for six SAF and six SAR resources being dedicated to Subnetwork B. The further distinction between Partition 1 and Partition 2 is that Partition 1 includes a mix of manned and unmanned missions including Freedom and STS on Subnetwork B. Partition 2 assumes that all manned missions are partitioned on Subnetwork A and all

unmanned missions are partitioned on Subnetwork B. . For each of the Subnetwork A resource cases, two alternatives are examined for the set of resources that demand access requests can access: both SA and MA and MA only. The next level of parameter variation is the demand access request type. Two cases are considered: all demand access requests are blocking and all demand access requests are window. The last parameter category is traffic load. For each of the cases defined by the tree diagram, we examined the third quarter 1998 baseline load (100 percent), 200 percent of the baseline, and 300 percent of the baseline. The additional traffic load was defined by adding a second set and third set of missions identical to those of the baseline traffic model with the spacecraft orbits off-set by 15 and 30 minutes, respectively, for the two additional cases. In addition to the cases defined by the tree diagram in Figure E-1, we also examined the impact to blocking probability for smaller window size and longer contact time for each demand access request.

E.3 Model Results

E.3.1 Partitions Results

Three cases were examined for traffic partitioned between Subnetwork A and Subnetwork B. The baseline case assumes that the missions listed in Table E-1 are partitioned onto Subnetwork B. The results in Table E-2 show a comparison of blocking percentage for the baseline partition with all traffic scheduled and the two partitions examined for demand access and scheduled traffic. The partitions were described in the previous section. Two demand access types were examined: blocking, or immediate, requests and window requests (i.e., requests that can wait if a resource is not immediately available). The blocking probabilities for the scheduled traffic were determined by the NPAS model. The values given for the scheduled traffic are the maximums of all missions using scheduled access. The maximum acceptable blocking probability is 10 percent.

The results show that moving Freedom and STS onto Subnetwork A reduces blocking on Subnetwork B. Both Freedom and STS require full coverage during their TDRSS visibility times from single access resources. When these two sets of demands are moved to Subnetwork A partition, contention on the remaining resources on Subnetwork B is reduced on the remaining single access resources.

This result occurs even though two more SAF and two more SAR resources are dedicated to Subnetwork A users.

Another conclusion drawn from the results is that demand access cannot provide an acceptable level of blocking if the missions cannot wait if a resource is busy when a request is made. This conclusion holds true if the current resource allocation between SA and MA resources is maintained. Also, the blocking probability can be acceptable if requestors have the flexibility to wait for a resource if it is busy.

E.3.2 Demand Access Results

The model was used to investigate resource allocation, workload growth, coverage requirements, and window size issues for demand access requests.

Table E-3 shows the percentage of demand access requests blocked for the baseline traffic load, two times the baseline load, and three times the load for Scenario 1. Blocking probabilities are given for all blocking, or immediate, requests and all window requests. In this case demand access transactions can access both SA and MA ATDRSS resources. The results show that if all requests are the blocking type, the percentage of requests blocked for the baseline load is unacceptable at 30 percent. However, if the requests are all window types, where the window is 100 percent of a spacecraft's visibility time, the percentage blocked is zero for the baseline case. The results for the window case shows that the percentage is still relatively low at eight percent for 200 percent of the baseline load, and slightly above acceptable at 12 percent for 300 percent of the baseline.

A closer look at the results in the previous table is given in Table E-4. The results in these tables show blocking percentage broken down by SA and MA resources for all blocking requests and all window requests. Both cases show that the blocking is occurring on the SA

resources with almost no blocking on the MA resources even at 300 percent of the baseline load. These results suggest that it is highly desirable to move the demand access requests to the MA resources assuming that mission requirements can be met in terms of bandwidth and any other transmission characteristics. The results for the cases where all demand access requests are served by MA resources only are described below.

When all demand access requests are moved to multi-access resources the percentage of requests blocked for both blocking and window types is below 10 percent for the range of traffic examined (Table E-5). If all demand access requests are window type, then less than one percent of the requests are blocked at 300 percent of the baseline. The blocking percentage for blocking request type is less than one percent for the baseline load and slightly less than 10 percent for 300 percent of the baseline.

The results of doubling the mission contact times are summarized in Table E-6. The table shows the blocking percentage for 100 and 200 percent of the baseline contact times for two cases: demand access requests use both SA and MA requests or use only MA resources. For the former case blocking percentage increases significantly for both request types. There is little or no change in the blocking percentage where all demand access requests are served by MA resources only. These last results also show the robustness of allocating single access resources to scheduled traffic and MA resources to demand access traffic. The NPAS model results show that all scheduled missions have a blocking percentage of less than five percent. These results show that the blocking percentage for demand access requests is less than five percent even if the workload requirements are much greater than the baseline.

The results for Partition 2 are given in Tables E-7 and E-8, respectively, for demand access traffic on both SA and MA resources and on MA resources only. The results follow the same trends as the previous results. Having the flexibility of a window reduces blocking percentage as does moving demand access requests to multi-access resources.

E.3.3 Window Size Results

The impact of window size was examined by varying the window size. The results, given in Tables E-9, show the change in blocking percentage and waiting time. These results are based on Partition 1. The results show that the blocking percentage decreases from 30.6 percent down to zero as the window size increases from zero to the full visibility time. If the window size is reduced to 50 percent of the total visibility time, blocking is reduced to an acceptable level at 5.6 percent with a waiting time of about five minutes for requestors that have to wait. Around 12 percent of requests have to wait for the 50 percent window size case. The waiting time increases as the window size increases but is less than ten percent of a typical orbit time for the maximum window size. These results assume that both SA and MA resources are used by demand access window requests. Reducing the window size has no impact on the case where all demand access requests are served by MA resources only since there is no blocking for that case.

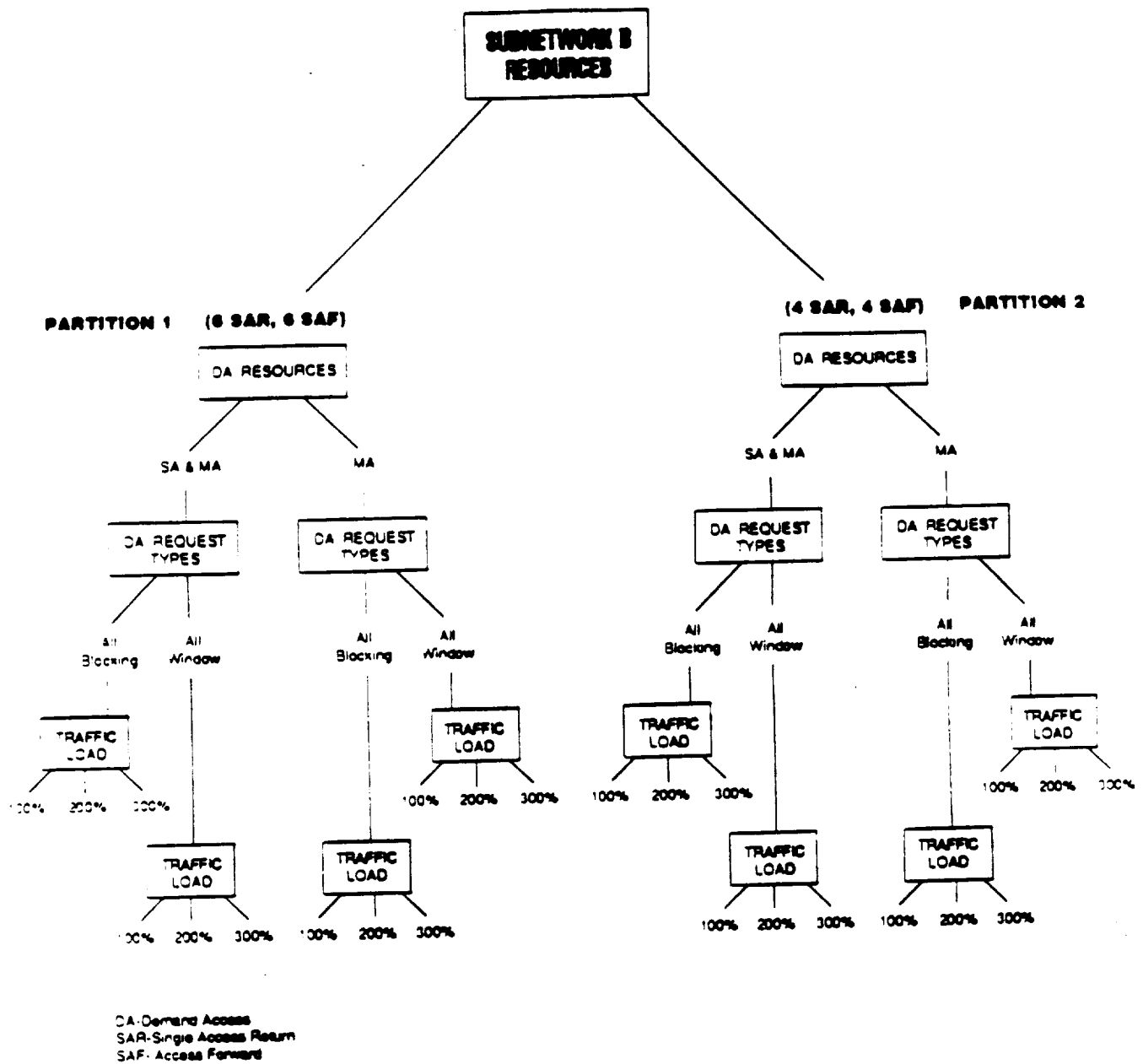


Figure E-1: Evaluation Scenarios

DEMAND ACCESS MISSION	SAF	SAR	MAF	MAR	CONTACTS/ORBIT
GRO	30 20	30 20			1/2 2 PER DAY
AXAF	20	20	10	20	1/2 1
TRYM		10			1
LYMAN	15	15			1
EOS	30	30		100%	1 1
SCHEDULED MISSION	SAF	SAR	MAF	MAR	CONTACTS/ORBIT
FREEDOM	100%	100%			1
STS	100%	100%			1
HST	15	20 20 60			1 3/day 1/week for 4 orbits
			10	100%	1 1
OSL	480 5	480 5			1/day 1/40 min
			20	100%	1 1
GP-B1			15	100%	1/day 1
COLDSTAT	10	10		100%	1 1

NOTES:

1. Times are in minutes
2. Abbreviations:
 - SAF - single access forward (K & S bands)
 - SAR - single access return (K & S bands)
 - MAF - multi-access forward
 - MAR - multi-access return

Table E-1: Baseline Traffic Requirements

<u>Partition</u>	<u>Scheduled</u>	<u>Demand Blocking</u>	<u>Access Window</u>
Baseline	5.0	-	-
Partition 1	5.0	30.6	0.0
Partition 2	0.0	22.6	2.7

Table E-2: Partitions Sensitivity - Blocking Percentage

<u>REQUEST TYPE</u>	<u>PERCENT OF</u> <u>100%</u>	<u>BASELINE</u> <u>200%</u>	<u>WORKLOAD</u> <u>300%</u>
All Blocking	30.6	36.4	40.5
All Window *	0.0	8.4	12.3

Table E-3: Partition 1 - Blocking Percentage for Demand Access Requests on SA and MA Resources

<u>REQUEST TYPE</u>	<u>RESOURCE TYPE</u>	<u>PERCENT OF</u> <u>100%</u>	<u>BASELINE</u> <u>200%</u>	<u>WORKLOAD</u> <u>300%</u>
All Blocking	SA	44.9	55.6	59.2
	MA	0.0	0.0	0.9
All Window *	SA	0.0	12.4	18.0
	MA	0.0	0.0	0.2

Table E-4: Partition 1 - Blocking by Resource Type

<u>REQUEST TYPE</u>	<u>PERCENT OF</u> <u>100%</u>	<u>BASELINE</u> <u>200%</u>	<u>WORKLOAD</u> <u>300%</u>
All Blocking	0.8	3.4	9.8
All Window *	0.0	0.1	0.6

Table E-5: Partition 1 - Blocking Percentage for Demand Access Requests on MA Resources Only

Notes:
 * Window size equal to 100 percent of TDRSS visibility time

<u>REQUEST TYPE</u>	<u>SA & MA</u> <u>CONTACT TIME</u>		<u>MA ONLY</u> <u>CONTACT TIME</u>	
	<u>100%</u>	<u>200%</u>	<u>100%</u>	<u>200%</u>
All Blocking	30.6	48.1	0.8	3.8
All Window *	0.0	22.0	0.0	0.0

Table E-6: TDRSS Contact Time Sensitivity - Blocking Probability

<u>REQUEST TYPE</u>	<u>PERCENT OF BASELINE WORKLOAD</u>		
	<u>100%</u>	<u>200%</u>	<u>300%</u>
All Blocking	22.6	31.6	36.1
All Window *	2.7	5.8	9.9

Table E-7: Partition 2 - Blocking Probability for Demand Access Requests on SA and MA Resources

<u>REQUEST TYPE</u>	<u>PERCENT OF BASELINE WORKLOAD</u>		
	<u>100%</u>	<u>200%</u>	<u>300%</u>
All Blocking	0.1	3.3	7.9
All Window *	0.0	0.2	0.4

Table E-8: Partition 2 - Blocking Percentage for Demand Access Requests on MA Resources Only

<u>Window Size</u> **	<u>Blocking %</u>	<u>Waiting</u>	
		<u>Time (min)</u>	<u>% Waiting</u>
0	30.6	0	0
50	5.6	4.96	11.9
100	0.0	8.73	2.3

Table E-9: Window Size Sensitivity

Notes:

- * Window size equal to 100 percent TDRSS visibility time
- ** Percent of TDRSS visibility time

APPENDIX F

SNC/ISC Dataflow Analysis

This appendix documents the analysis used to evaluate the architectural alternatives described in section 3.3.1.1 (Real-time vs non real-time functions). Table F-1 shows the estimated frequency and time criticality of primary control messages. The term time critical is used to refer to messages that are used in a real-time (or near real-time) control transaction, e.g. fault event control or service parameter change requests during a service event. The dataflow analysis assumes that these transactions must be completed within few seconds, i.e., transit delays for individual messages should be of the order of 100 msec. Non time-critical messages are those that are used for routine control transactions, e.g., event accounting. These messages can be delivered in background mode.

The logical flow of the control messages between ATGT, CDOS, SNC/ISC system(s) and the User POCCs is shown in Figure F-1. The dataflow rates (in messages/sec) shown in the figure is for the worst case scenario, i.e. maximum control message traffic. In the worst case scenario will the ATDRSS constellation will be supporting 28 concurrent events (8 single access service and 20 multiple access service) and all events will be short duration (few minute) events. While this is highly unlikely, the SNC/ISC system must be designed to support peak event rate.

All architectural alternatives assume that the SNC/ISC system(s) will use a hierarchical control scheme based on the encapsulation concept. With this approach, the real-time SNC/ISC system receives summary results of monitoring information and exception reports. The service provider (ATGT, CDOS and NASCOM) control/management systems will be responsible for continuous monitoring (collection and processing) of fault and performance information. Therefore, the frequency of time critical message traffic to/from the real-time SNC/ISC system will be at most few messages per second per event. Therefore, in the worst case scenario, the number of monitoring messages per second is estimated to be approximately 100/second (28 concurrent events, 3-4 messages per event).

The real-time SNC/ISC system will process the performance monitoring information from different service providers and send a summary progress report to the User System every 5 seconds (this frequency was assumed to match the current practice). In addition, the ISC will also process fault monitoring information and if a fault event occurs, it will send the necessary information to the affected User System. Therefore, the total frequency of messages between the real-time SNC/ISC and a specific User System is estimated to be less than one/second.

The non real-time SNC/ISC system deals with resource availability, scheduling, accounting and security management. The only time critical messages received/sent by this system deal with on-demand scheduling. In the worst case scenario, the number of such messages is estimated to be few per minute. The number of pre-planned service requests and associated shift negotiation messages (for a particular User System) is estimated to be less than one per minute. The number of accounting and security management messages is estimated to be few per event. Therefore, the aggregate message traffic between the non real-time SNC/ISC system and a specific User System is estimated to be less than one/second.

The peak rate for security and accounting messages (28 concurrent short events of few minutes each) between the non real-time SNC/ISC system and ATGT is estimated to be approximately one per second (10 messages per event, 280 messages in five minutes).

The non real-time SNC system will prepare a composite schedule and provide it to all service providers and the real-time SNC/ISC system. Currently this is 3 times per day (once every 8 hours). In the ATGT era, the frequency may be increased to once per hour. The same is true for the resource availability schedule provided by various service providers, including ATGT. Therefore, the worst case aggregate dataflow between the non real-time SNC/ISC system and a service provider (such as ATGT) is estimated to be two file transfers/hour.

PRIMARY CONTROL MESSAGES

Estimated Peak Rate

I. Time critical messages

Ia.	Fault monitoring information (includes resource monitoring)	few/sec
Ib.	Performance monitoring information	few/event
Ic.	On-demand service/parameter set-up requests	few/minute
Id.	Event progress reports	12/minute
Ie.	Fault event control (during an event)	few/event

II. Non Time critical messages

IIa.	Pre-planned service request response /shift requests	1000/day
IIb.	Pre-planned scheduled file	one /hr
IIc.	Accounting information	few/event
IId.	Security information	few/event
IIe.	Resource availability plan file	one /hr

Table F-1

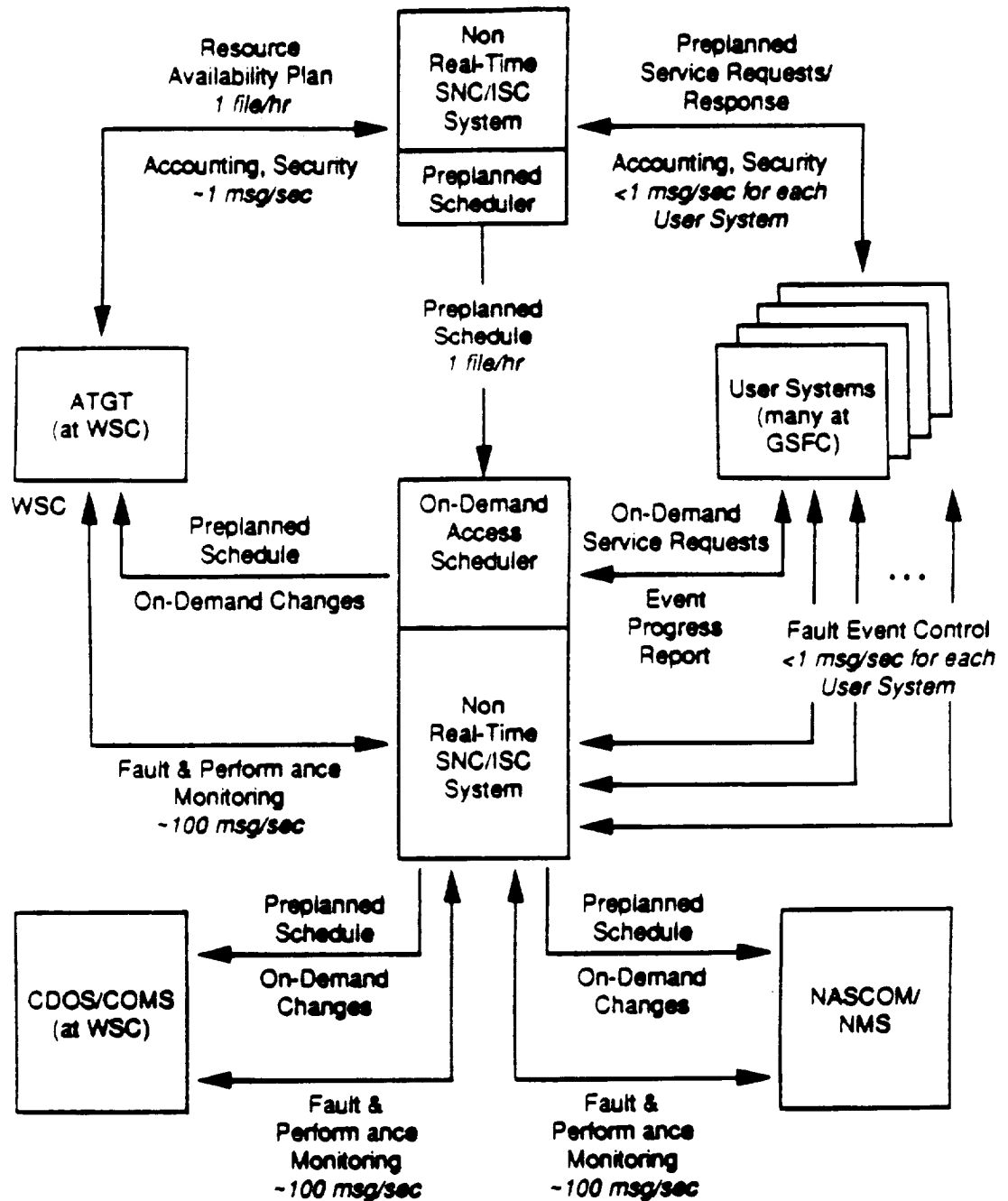


Figure F-1. Logical Flow of Control Messages

The exact content, size and frequency of the control messages will be specified during the system design and implementation phase. However, it is our estimate that generally the message size will be of the order of 10^{**3} bits (125 octets). Therefore, the required sustained throughput for the time critical dataflow paths is estimated to be 10^{**5} bits per second (10^{**3} bits/message times 10^{**2} messages per second), i.e., 100 Kbps.

A sustained throughput rate of 100 Kbps for short messages can be easily supported over a redundant LAN or over redundant long-haul reliable links (e.g. fractional T1 links with TP4 protocol). The cost for long-haul communication are higher. However, in the overall scheme of things, the cost of a long-haul fractional T1 circuit is not very significant compared to the total operational cost of the SNC/ISC system. Therefore, collocating the ATGT, CDOS/COMS, and the real-time SNC/ISC systems at WSC would be provide marginal cost savings compared to locating the real-time SNC/ISC system at GSFC. On the other hand locating the non real-time SNC/ISC system at GSFC would facilitate more effective communication with User Systems at GSFC, without any significant cost impact.



Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle Advanced System Engineering and Network Planning Support - Final Report 'A' Team Study		5. Report Date December 1990
		6. Performing Organization Code
7. Author(s) David H. Walters, Larry K. Barrett, Ronald Boyd, Suresh Bazaj, Lionel Mitchell, Fred Brosi		8. Performing Organization Report No.
9. Performing Organization Name and Address CTA INCORPORATED 6116 Executive Boulevard Rockville, MD 20852		10. Work Unit No.
12. Sponsoring Agency Name and Address National Aeronautics and Space Administration Washington, DC 20546-0001 Goddard Space Flight Center		11. Contract or Grant No. NASS-30680
		13. Type of Report and Period Covered Final 8/20/90-12/21/90
		14. Sponsoring Agency Code
15. Supplementary Notes		
16. Abstract <p>The objective of this task was to take a fresh look at the NASA Space Network Control (SNC) element for the Advanced Tracking and Data Relay Satellite System (ATDRSS) such that it can be made more efficient and responsive to the user by introducing new concepts and technologies appropriate for the 1997 timeframe. In particular, it was desired to investigate the technologies and concepts employed in similar systems that may be applicable to the SNC.</p> <p>The recommendations resulting from this study include resource partitioning, on-line access to subsets of the SN schedule, fluid scheduling, increased use of demand access on the MA service, automating Inter-System Control functions using monitor by exception, increase automation for distributed data management and distributed work management, viewing SN operational control in terms of the OSI Management framework, and the introduction of automated interface management.</p>		
17. Key Words (Suggested by Author(s)) Space Network Control, Resource Allocation, Scheduling, Network Management, TDRSS		18. Distribution Statement Unclassified Unlimited
19. Security Classif. of this report Unclassified	20. Security Classif. of this page Unclassified	21. No. of pages 136
		22. Price